

Agenda for a US president

The outcome of the election in the United States will matter greatly both for the indigenous research enterprise and for its international ramifications.

JOURNALS such as this do not take sides in election campaigns, and for good reason. Rarely does science and a prospective government's policy on research dominate the hustings, which is on balance fortunate. In the long run, it is better for the research enterprise that its importance as an engine of constructive change should be widely appreciated, by all contenders, than that there should be single-issue elections centred on the treatment of research. But that does not mean that journals whose chief interest is the welfare of the research enterprise are indifferent to the outcome of elections. This week's presidential election in the United States will have been especially influential.

There have of course been exceptions to the rule that research plays no part in elections. In Britain in 1964, the then Mr Harold Wilson was elected prime minister on the strength of a promise that "white-hot technology" would make Britain prosperous; even a cursory appraisal of his plans would have shown that they would end in tears. There was another opportunity in 1980, when M. Francois Mitterrand was elected president of France on a research ticket whose benefits have been more substantial. This year's election campaign in the United States, which ended on Tuesday this week, has predictably had little to say about science and research since the exit of Mr Paul Tsongas from the contest in the summer. Mr Bill Clinton, the Democrat, has offered to put flesh on the bones of President George Bush's declaration four years ago that he would be the "education president", which is a step in the right direction. Mr Ross Perot has been talking about the need that the United States should build up a greater degree of technological skill. But research as such has not been a central issue.

Yet the outcome of the election, which will be known by the date of this issue, will matter a great deal for research everywhere. For the past half-century, the United States has been the world's chief source of intellectual innovation and technique in most fields of science and technology. That many of those who have individually contributed to this imaginative effort were not born in the United States is not a chauvinistic shame, but the opposite — a measure of the traditional readiness of people in the United States to respect the skills of others and to enable them to be put to good use. By that means, the United States has won both profit and honour for itself, and has helped to sustain the research enterprise elsewhere.

But now the recipe has two serious flaws. By the impoverishment of public education, too few indigenous Ameri-

cans participate in advanced education, which acts against industrial change and is also socially divisive. And US administrations, helped or even egged on by the Congress, have come to endow research with chauvinistic attributes. The attempt to persuade Japan to contribute a fifth of the cost of the Superconducting Super Collider is humiliating for both parties, the general paranoia in the United States about Japan's success in high-technology manufacturing is a dangerous misreading of the causes of the relative US decline in such fields. Whichever candidate has won, it is to be hoped that he will give these issues the attention they urgently deserve. But how?

Populist Perot, the most numerate of last week's candidates, seems to have grasped one obvious need. Although the drumbeat of his complaint against the size of the federal budget deficit seems to have been born of a homespun (and old-fashioned) belief that even governments should balance their books, the federal deficit has impeded the management of the US economy for the past several years. The newly elected president will be able to meet the electorate's demand for more jobs only by making room for extra public spending, which will entail extra taxes and reduced transfer payments, "entitlements" as they are called. Like it or not, the new incumbent at the White House will have to be a "tax and save" president, to adapt Bush's election jibe against Clinton. Let us hope that he does not choose to "save" on research and that he learns to spend on public education — or make state governments spend.

The entitlement issue is potentially more seriously divisive. The growing armies of the elderly and of the poor have a proper stake in whatever prosperity the United States can muster in the years ahead. The new man's problem is that the constituency of the poor is a novel and even un-American constituency (the long-standing poverty of Appalachia notwithstanding). Traditionally, successive waves of immigrants have been sustained through poverty by the characteristic optimism of the United States. That is also the chief source of the institutional and technical innovation that made the United States a formidable industrial power. But this election has shown the United States to be almost despondent at the social and economic problems it confronts. The new president's most important task is to change that mood. (That, in his odd, is what Perot has been saying.) Unless the new president can work that trick, the ebullience of US research since the Second World War is not the only virtue at risk of dissipation. □



Eppur si non muove

The Vatican's half-hearted rehabilitation of Galileo will not prevent the recurrence of errors of the same kind.

So Galileo is in the clear, then. That is what the Vatican said last week. More than a third of a millennium after the grand old man was compelled by the Inquisition to disavow his heliocentric hypothesis of the Solar System, it is now acknowledged that he should not have been put to that shameful test. But ungenerously. The Vatican's twelve-year inquiry into the Galileo business concludes that both Galileo and his inquisitors acted "in good faith", observing that Galileo's case would have been stronger if his evidence that the Earth goes round the Sun had not been "inconclusive". The basis for that opinion is that the Inquisition, faced with the account of the Creation in *Genesis* on the one hand and Galileo's chain of inference on the other, had no choice but to believe the former. That is a little like saying that, if there had ever been a confrontation over Darwin's theory of evolution by natural selection, *Genesis* would have again been taken, literally, as gospel. A century ago, of course, the Church had learned to be more circumspect.

With the passage of so much time, many will say that this old tale, however heroic, has no present significance. But that would be mistaken. The Galileo business remains a present problem because it provides a perpetual licence for prejudice in the evaluation of discovery. Card-carrying creationists may now be thinner on the ground than in the seventeenth century, but the damage they can do, especially among the young (not to mention impressionable school boards in the United States), is considerable. More seriously because more generally, the Galileo business has endowed with a measure of gentility the doctrine that the process of discovery is properly assayed by the banal test of common sense. (The other side of that coin is the tendency of those with unpublished and unpublishable theories to plead Galileo in their aid.)

The irony in the Vatican's decision last week is that the test of common sense does not apply. Does the Earth go round the Sun (or, more accurately, around the instantaneous centre of gravity of the Solar System) in any but a relative sense? Prescient Galileo was probably too good a scientist to have committed himself to an absolute view on such a question, but such prudence would hardly have seemed convincing to the Inquisition. That is the great misfortune buried in last week's announcement from the Vatican. It further reinforces the general belief that discoveries should be cut and dried — black replaced by white, or *vice versa*. But that, of course, is not the way of the world. Karl Popper is not the sole authority on that point. There is also the breathtaking record of discovery that Galileo's sense of physics helped to usher in. The Vatican owed him a less grudging recantation. It has not moved much in 359 years. □

Activism's toll

The US Congress must reverse its growing tendency to play the role of peer reviewer.

THE US Congress, as usual, has made a muddle of its new science budget. It has managed to be simultaneously stingy in funding the research that the science agencies — and by extension the researchers they support — have asked for, and generous in funding that research that it itself finds politically expedient.

Nowhere is this more true than in the defence budget, where Congress buried (along with at least \$200 million in other research earmarks) \$210 million for breast cancer research. Thanks to a lobbying campaign massive even by Washington standards, legislators decided to win a few votes with the women back home and voted last month for one of the largest one-year increases of a research program in history. Why the defence budget? It allowed legislators to get around their own caps on non-defence spending.

When this sort of thing takes the form of a water project or a laboratory in somebody's home district, it is called "pork," and is rightly ridiculed. But when it goes to research directed at a specific ailment, it is called "disease earmarking." By any name, it has become the hottest game in town. This year's pork reached epic proportion. The situation has become so bad that the pig farmers have started to take out newspaper advertisements complaining that their product is being slandered by the comparison.

Inspired by the successes of the AIDS activists, new advocacy groups are lobbying for more research funds for everything from Lyme's disease to Alzheimer's. And Congress seems unable to resist their pitch. This year's budget set a record for disease earmarks — more than \$400 million dollars in total. Appropriations bills now read like medical dictionaries: Osteoporosis. Lupus. Rheumatoid arthritis. Hypertension. On and on, page after page.

Take the breast cancer earmark. Although the Pentagon's first inclination was to spend its windfall on mammography machines at Army bases, it now looks like most of the money is actually going to go to research. But it is hard to say just where the government is going to find enough breast cancer scientists — and enough good research projects — to more than double the current effort overnight.

Even researchers — those who stand to benefit most from this congressional largess — are fuming over the crass politics of it all. And as Congress and the disease groups eat up more and more of the pie, research projects that rate highly on their scientific merits are coming in a weak second.

The message is clear: peer review is out and activism is in. Activists beware: next time your legislator offers to send some research money your way, watch whose pocket it is coming from. Looting the defence budget may work this year, but it will not be long before the disease-of-the-month crowd starts feeding on itself. Taking from one scientist to pay another to get votes is not medical progress — and does not mean that it is money well spent. □

Synchrotron beamlines open as plans are laid for repairs

Grenoble. Shutters are being opened this week to allow X-rays through for experiments at the European Synchrotron Radiation Facility (ESRF) here despite continuing problems with the ECU450 million (US\$600 million) facility's concrete floor. Later this month, officials are expected to approve rebuilding more than half of the floor, at a cost of ECU2 million, to correct the most serious problems with the least disruption to the scientific programme.

The two beamlines coming alive this week are for crystallography and work on tiny samples of rare materials or at high pressure, with a third, multipurpose beamline opening later this month. The lines will allow researchers to work with high brilliance, the facility's special feature, and to learn more about such technical problems as releasing large amounts of heat. Scientists also want to know whether the grouting injected into the floor underneath these beamlines will reduce vibration adequately. "We will be getting photons", says one scientist, "but whether we can handle them is another matter".

The trouble with the floor has its roots in a decision, made largely for reasons of politics and economics, to build the synchrotron at Grenoble rather than at a competing site at Strasbourg despite its being subject to seismic vibrations and city noises, its proximity to rivers and roads and its lying on postglacial alluvial terrain. Comparable machines rest on rock, with tidal movement in some cases being the only detectable vibration.

The problems were compounded during construction, starting in 1989, by soil that was not compacted, concrete slabs that were too thin and surfaces that were worked prematurely. In addition, a steel reinforcement grid was laid below rather than within these slabs. "A supermarket floor", scoffs one scientist, referring to a type of project commonly handled by Bouygues, the large French civil engineering consortium that performed the work.

Voids below the surface began to appear in June 1991, followed by broken corners, disintegrating surfaces and misalignment of the concrete slabs. Walking on the edge of a slab was enough to cause a vibration 200 times greater than the highest levels acceptable. Last month, as part of litigation between Bouygues and ESRF, a French court appointed two experts to study the case and determine who should bear the cost of repairs.

On 24 November, the ESRF Council will hear recommendations from two advisory bodies that it should replace one major sector of the 250-metre-diameter floor while

allowing experiments to continue on the opposite side. Under that plan, repairs would begin in spring 1993 and take 20 weeks, with the entire ring being shut for six or



The alluvial terrain appears below the compacted gravel, sand and concrete beneath the floor of the ESRF; a pen indicates the thickness of the slab.

seven weeks. The leading design, from engineers at Britain's Harwell Laboratory, would involve digging up the existing 15

cm concrete floor and installing a layer 70 to 80 cm thick. Workers compact the natural gravel bed first, and then put down thicker vibration-absorbing layers before laying concrete.

It is generally agreed that the problem cannot be eliminated because of constraints imposed by existing buildings, principally the ring. A better design would have been to put the whole installation on a single reinforced foundation rather than having the ring and experimental hall on separate concrete platforms.

The cost of repairs amounts to about 10 per cent of the facility's annual budget, and a significant part of the money available for research. Even so, the spirits of ESRF scientists have risen with the arrival of Yves Petroff, who in January succeeds Ruprecht Haensel as director-general. Outside research teams have been offered 48-hour slots for simple demonstration projects to run in parallel with the commissioning of the beams, and officials are confident that at least seven beamlines will be open to users by spring 1994.

"People have been waiting a long time for a facility like this", says one scientist. "There is nothing else like it in the world."

Susan Biggin

French modify polio plant in India

New Delhi. A French vaccine manufacturer has agreed to retool a plant under construction outside Delhi to make the type of polio vaccine preferred by the Indian government. The decision reverses a position taken earlier this year (see *Nature* 356, 94; 1992) by the Institut Marieux that it would withdraw from the joint venture and stop work on the plant, expected to be ready by the end of the year. "We are putting good customer relations ahead of financial considerations", says a company spokesman in Delhi.

The change of heart appears to be the result of talks last month in Paris between science ministers of the two countries and the subsequent visit of the Indian prime minister, P.V. Narasimha Rao, to France. Rao was apparently able to convince the French president, Francois Mitterrand, that the company's withdrawal would jeopardize India's immunization efforts.

In 1989, the Institut Marieux joined with the Indian government to start building the biggest vaccine plant outside Europe for the manufacture of injectable polio vaccine. The company was persuaded to share its

technology because of the large Indian market and the likelihood that other developing nations might also purchase the vaccine.

But the project ran into trouble last year when India, following the recommendation of a report by the World Health Organization, decided to use the cheaper oral polio vaccine instead of the injectable vaccine that Institut Marieux planned to make. The company said that it would pull out of the project until it was allowed to market the injectable vaccine to India, its only assured customer. Several rounds of talks ended in a deadlock last summer and in July the company's chief executive office, Alain Audubert, said that "we have decided to freeze all further commitments to the project".

That is no longer the case. Later this month, a committee of experts from Institut Marieux and the Indian Department of Biotechnology is expected to submit a report on ways to convert the manufacturing plant, and company officials say that they are already working "to make up for lost time". If all goes well, the first batch of vaccine could be ready by the middle of 1993. **K.S. Jayaraman**

Opponents of US earmarks propose reviews to temper worst elements of growing practice

Washington. This year, US researchers are once again reacting with a mixture of shock and distaste to the new federal science budget. The US Congress has again riddled its appropriations with politically motivated earmarks (targeted projects), steering hundreds of millions of dollars towards research that scientists would never request — and may not even know what to do with now that they have been given it.

Conceding that they will never win the battle against such spending abuses, some powerful legislators are hoping at least to minimize the damage. They have started requiring earmarked projects to undergo some type of peer review to screen out the most egregious misuses of federal funding.

Legislators have been forced to take such a politically unpopular step by a budget process that many believe is out of control. The budget for fiscal year 1993, which began last month, contains \$400 million more for research on specific diseases than the amount federal science agencies requested. And academic 'pork' — appropriations to specified universities and research institutions — appears to have set a record as well.

Although the total is not yet known, the 1993 defence budget contains at least \$400 million, and the energy and water appropriations another \$115 million; in all, this year's earmarks appear likely to exceed last year's record of \$707 million, spread across 500 projects at 170 institutions.

This is despite last-minute heroics by Representative George Brown (Democrat, California), who succeeded, at some considerable political cost, in removing nearly \$100 million in earmarks from an energy bill shortly before its passage. But his victory was short-lived; each appears in a bill for military expenditure passed a little more than a week later.

Researchers struggling for more funding might forgive the congressional earmarking as a means to an end. If Congress wants to spend more on science, by whatever mechanism, why complain?

In fact, if the practice of earmarking really did mean a larger science budget, researchers would probably remain silent. But it does not. The 1993 budget for the National Institutes of Health (NIH), for example, contains at least \$183 million more

than NIH requested for research on such diseases as Alzheimer's and prostate cancer. Yet the overall NIH budget will grow by less than the rate of inflation, and the money to pay for what Congress has stipulated will have to come from other, less politically popular areas within the NIH budget, including basic research on cellular mechanisms and biochemistry.

Bernadine Healy, the NIH director, complained at a congressional hearing in

satisfy most language in its appropriations bills without abandoning its scientific priorities. But this year's budget is threatening even that practice. In an attempt to get around budget rules that limit the amount of spending on domestic programmes, legislators have now begun attaching disease research earmarks to the defence appropriations bill. Most prominent of these is \$210 million for breast cancer research in this year's budget (see *Nature* 359, 471; 1992). To maintain the illusion, Congress put the military in charge, asking only that NIH participate in discussions on ways to spend this money.

Photo by J. Allen Hensley

After rumours spread that the money would be used to pay for mammography machines and other hardware at Army bases, defence officials (under pressure from breast cancer groups) agreed to work with NIH to devise a research plan and review grant applications. But the Pentagon is still clearly in charge, which makes researchers uneasy.

Breast cancer is not the only battlefield. Inspired by the success of the AIDS activist groups in getting vast increases in research fund-

ing, a host of other groups have emerged in the past few years advocating more money for research on dozens of ailments, from Lyme's disease to leukodystrophy. And unlike some of the more established disease groups, the new coalitions are bypassing the public and heading straight to Congress.

This strategy pits one disease group — and its associated researchers — against another in a lobbying war for research funds. "It'll be a disaster if the research communities start attacking each other to get more of the money for themselves", says Harmon Eyre of the American Society of Clinical Oncology. In an attempt to prevent that, the society is encouraging its members to stick to the party line: more money for NIH overall will benefit everyone.

NIH officials are working on their own coping mechanisms. One option is to calculate each year the amount of money being spent on each disease, and publish these numbers in the agency's annual budget request. That would at least tell advocacy groups what NIH is already doing, says Jack Mahoney, NIH's chief financial officer. Perhaps the groups, seeing that their

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

More than 22,000 people participated in this year's 'Race for a Cure' for breast cancer in Washington DC.

September that it is "extraordinarily frustrating when we have congressional language that says we must spend \$250 million on breast cancer, and there is no money [provided in the budget] to do that".

The congressional budget process aggravates the situation. Most NIH earmarks appear in the Senate version of the appropriations bill, which typically supports a larger budget for NIH. The final budget, reached after negotiations between the two houses of Congress, is usually smaller, but the earmarks remain and thus take up more of the overall budget.

NIH at times can avoid the problem by redefinition, for example classifying a basic biology project as "breast cancer research". But university scientists must know enough so to describe their proposed project when they apply for a grant. And it is difficult for a researcher to anticipate which disease is going to be in fashion in the next round (although breast cancer seems a certainty for a few years). All too often, NIH must engage in creative accounting to assemble a research portfolio that will pass muster with Congress.

NIH has used such tricks in the past to

interests are shared by NIH, might then decide to lobby for the agency's entire budget. Yet even this plan may backfire; the groups may instead use the current NIH figure as a base and simply lobby for a hefty increase on top of that.

In the meantime, Congress is attacking the problem by looking at its own practices. Its first priority is controlling pork, but it also wants to see the money spent wisely. Brown has proposed several changes to the rules under which Congress operates that would make the most flagrant pork easier to kill, including an opportunity for legislators to raise objections during the appropriations process and the inclusion of a 'pork watchdog' in appropriations conferences. He and others have promised to give pork a run for its money next year, assuming that voters reelect them this week.

Another example is a \$20-million earmark for a AIDS vaccine trial in the Senate defence bill. In previous years, this language, which is intended to apply only to a gp160 vaccine made by MicroGeneSys, Inc. of Meriden, Connecticut, might have simply been approved as unaltered pork. But this year, in a modest concession to the scientific process, legislators decided to allow the NIH director, the director of the Food and Drug Administration and the secretary of defence to review the expenditure and spend the money elsewhere if they do not think that the trial should proceed.

It is not a coincidence that one of the authors of the amendment is Senator Sam Nunn (Democrat, Georgia), the chairman of the Armed Forces Committee and traditionally one of the strongest voices against earmarking. Although Nunn had no more success than Brown in removing pork from this year's defence budget, he was able to moderate it with a technique he hopes to repeat. Some \$75 million worth of earmarked projects cannot be released until they have passed a merit-review process based on the potential contribution each project would make "to the national scientific and technical posture".

Unfortunately, even that language has become controversial. In deference to Senator Daniel Inouye (Democrat, Hawaii), a member of his committee, Nunn stipulated that the reviewers should be from institutions that belong to the National Association of State Universities and Land Grant Colleges or the American Association of State Colleges and Universities. Although that group includes the University of Hawaii, in Inouye's home town, it excludes all 58 members of the Association of American Universities, which includes most of the largest US research universities.

Nevertheless, most legislators believe that some sort of merit review is better than none at all. And researchers are hoping that the money for earmarking, even if it is here to stay, will go towards the best science.

Christopher Anderson

Japan wants global guidelines for joint research projects

Tokyo, Washington & Munich. In another example of 'technoglobalism', Japan's Ministry of International Trade and Industry (MITI) is trying to persuade the world's leading nations to agree on guidelines for international industrial research and development projects. But the proposal, put forward at a meeting of the Organisation of Economic Cooperation and Development (OECD) in Paris on 19–20 October, has left Western nations perplexed.

As a counterweight to moves by the United States to protect its industrial intellectual property rights — known in Japan as 'technonationalism' — MITI has tried in recent years to encourage advanced nations to establish international industrial research and development projects. One example is the Intelligent Manufacturing Systems (IMS) project to develop the automated factories of the future, which MITI unveiled in 1990 to an unsuspecting world (see *Nature* **343**, 496; 1990).

The project is an attempt to direct the resources of the world's most advanced companies and research laboratories towards sophisticated computerized manufacturing systems, but it met stiff opposition from the United States and Europe on such issues as intellectual property rights. A feasibility study is finally under way after two years of negotiations (see *Nature* **355**, 755; 1992).

The difficulties with IMS encouraged MITI to make its latest proposal. "The guidelines/checklists could provide model procedures to start an international project [with] example agreements to be adopted and model arrangements for intellectual property rights" says the proposal to the OECD's Committee on Science and Technology Policy.

Thomas Ratchford, associate director for international affairs at the White House Office of Science and Technology Policy and part of the US delegation at the OECD meeting, says the Japanese proposal "is not very carefully defined". European delegates agree that the proposal is vague, and Japan has been asked to revise it for a meeting this winter. "We agree in principle with their approach", says William Booher of the Technology Administration within the US Commerce Department, "but there were real concerns about some of the substance".

One problem raised by the Western nations is the idea of applying guidelines to what are essentially unique ventures. "What the Japanese seem to want, for example", says Ratchford, "is a standard IPR [intellectual property rights] clause that can be pulled off the shelf and inserted into a new technology agreement. But it does not work that way."

Another problem arises from the differ-

ing relationships between government and industry in each country. Whereas MITI launches national projects to encourage Japanese companies to develop commercial technologies, the US government adopts much more of a 'hands-off' approach (except perhaps in the area of defence), leaving companies to form their own collaborations.

Masaya Yasui, deputy director of MITI's technology policy research and analysis division at the Agency of Industrial Science and Technology, accepts that the proposal has caused "misunderstandings". "People fear the guidelines may have restrictive power", he says. "We do not want to stimulate such anxiety. However, it's easy to point out differences [between countries and projects]. Surely it's much more productive to find common principles".

Japan also wants to improve communications between nations so that the world learns about new technology projects as early as possible. For example, Korea is unhappy that it was not included in the IMS project, and Japan feels that it is not well informed about European-based international projects.

European delegates agree that it will be very useful if the Japanese proposal establishes rules for joint projects and creates a mechanism for explaining their purpose. According to an OECD spokesman, Japan encounters such problems every time it seeks international partners.

Japan's proposal was not the only one made at last month's OECD meeting. The United States suggested a study of government-supported programmes on critical technologies to move forward in parallel with the Japanese initiative. Both would be coordinated by a working group on technology policy first proposed 18 months ago by the United States. It is likely that the proposals would be linked with an effort already underway to establish guidelines for international big-science projects such as the US space station and the Superconducting Super Collider. The Japanese say that they are confused by the timing of the US proposal, which US officials see as a "logical next step" in proposing future joint projects.

All these efforts will take some time to bear fruit. The OECD is not inclined to act quickly; its preferred style is lengthy discussion, followed by an analysis of the issue culminating in a report that may form the basis for action. In addition, such activities require the OECD to solicit contributions from member countries to supplement its meagre resources, and making those financial arrangements also takes time.

David Swinbanks, Jeffrey Mervis & Alison Abbott

Dutch shake up funding council, shift its emphasis

London. The basic research council in the Netherlands plans to shift money from individuals to large projects involving many researchers as part of an enforced restructuring. The Netherlands Organization for Scientific Research (NWO) changed course last month after having failed to convince the Dutch government to spend more on science.

This year, NWO will award 470 million guilders (US\$280 million), around 10 per cent of government spending on research, to 3,700 researchers for projects and equipment at 13 universities and at its own research institutes. NWO's governing body, the general board, allocates money to six councils, which in turn distribute grants to foundations covering 34 disciplines. The restructuring follows a threat in NWO's five-year plan that if the government did not increase its budget it would pare down its fields of interest and shrink its structure to match its budget.

Ninety per cent of NWO's budget at present goes to individual researchers, many of them at the beginning of their careers. The plan is to lower this proportion to 50 per cent and give larger subsidies to small research groups. NWO expects the slack to be taken up by the Ministry of Education and Science, which last year began to fund graduate education directly.

The restructuring affects the 34 foundations, each representing a particular research interest, through which NWO funds are channelled. NWO intends to withdraw its backing from those foundations with annual budgets below 4 million guilders on the grounds that they are too small to influence the field and are costly to administer. As well as releasing funds, the changes are

expected to simplify the grants procedure, allowing two rounds a year rather than one.

While the foundations may continue to exist to improve communications among researchers, funding will devolve to the councils. Although none has been singled out, some disciplines will undoubtedly suffer.

Seventeen of the 34 foundations and departments will lose their NWO status by May next year. They include the computer science, earth sciences and biophysics foundations, the meteorology and physical oceanography working groups and all but one of the programmes in the humanities and social sciences councils. The mathematics foundation will also lose its NWO status but its institute, the Centre for Mathematics and Computer Science in Amsterdam, is expected to remain open. The government has yet to respond, but NWO is not expecting opposition.

Meanwhile, the recession is forcing the Ministry of Education and Science to look more closely at its research budget of 4.7 billion guilders. It wants to cut library spending by 12.8 million guilders and terminate several projects on the dissemination of information. The Ministry also wants to reduce the Dutch contribution to CERN, the European Laboratory of Particle Physics, hoping that the difference will be made up by non-subscriber countries. The education ministry awards a third of its budget directly to universities, with the rest going to researchers through NWO, the Royal Netherlands Academy of Science and Letters, the Netherlands Organization for Applied Scientific Research and a number of agricultural and medical institutes under the control of other ministries.

Ian Mundell

Max Planck institute gets new life

Munich. The Max Planck Institute for Astrophysics at Garching, rumoured earlier this year to be closing, has won a stay of execution that may become permanent. Its strongest sign of life is a renewed search for a new director by the Max Planck Society, which runs nearly 60 research institutes in Germany.

The society has for more than a year been seeking a replacement for Rudolf Kippenham, who retired as director last year (see *Nature* 358, 267; 1992). Controversy arose when the society insisted that Kippenham be succeeded by a single director even though most institutes have a group of up to five directors to share the burden.

Scientists in Garching have always argued that a rotating system would be much easier to implement.

Now they are having their way. The society has formed a small committee with an indeterminate deadline to investigate this and the possibility of a foreign director. With the boundaries so widened, says Garching scientist Peter Schneider, the danger of closure looks much reduced.

The Max Planck Society says that talk of closure was unfounded and that its new search for a director is independent of a groundswell of support for the institute from astrophysicists around the world and in the media.

Alison Abbott

US drug trials said to ignore gender differences

Washington. Pharmaceutical companies do a poor job analysing how their proposed drugs will affect women, according to a report last week to the US Congress by the General Accounting Office that lays part of the blame on the Food and Drug Administration (FDA). The report recommends that the FDA require companies to look more closely at gender differences and to enrol a minimum number of women in clinical trials of relevant drugs.

The report examined late-stage trials for 53 new drugs approved by FDA between January 1988 and June 1991. (Women of childbearing age may not participate in earlier phases of clinical trials.) GAO says that a third of the drugs were tested in trials that included fewer than 250 female subjects, the minimum that FDA believes necessary to detect gender-related differences in drug safety and efficacy. Eight of the 13 drugs for cardiovascular disease, a leading cause of death in women, were not tested on enough women, the report found.

Fewer than half the drugs were analysed by gender, although evidence suggests that women and men react differently to some drugs. Only 12 per cent of the drugs were analysed for their reaction with either female hormones or oral contraceptives.

The report says that FDA guidelines on including women are ambiguous: they recommend that all those who will take the drug be represented in clinical trials but do not mention women. In a 1991 survey by the Pharmaceutical Manufacturers Association, nearly half of its members said that the FDA did not specify that women should be included in clinical trials. Some of the blame also falls on industry: although the FDA emphasizes the importance of analysing data for gender differences, more than half the companies submitting new drug applications in 1992 failed to do so.

The FDA says that its reviewers will begin mentioning this requirement to companies at the start of the testing process. The FDA also agrees that drug companies should analyse the way in which their products interact with female hormones and oral contraceptives.

But the agency disputes the report's claim that many drug trials enrol too few women. It believes that the figure of 250 should be a guide rather than a floor; for some drugs, testing 100 women is sufficient. FDA officials say that another problem is how to reconcile the fact that women suffer from some diseases later in life than men with restrictions on the use of the elderly as test subjects.

Traci Watson

New biology institute in Milan prizes quality above loyalty in effort to become world-class

Milan. The future of biological research in Italy brightened last month with the opening of an ambitious new research institute in Milan that wants to stand outside the traditional Italian system of staff appointments



The Dibit Institute

and make its mark internationally.

The Dibit Institute cost more than L10 billion (US\$7.5 million) to build and was financed by the private foundation H. San Raffaele, which runs a large hospital and research centre in Milan. It has attracted some of the best biologists in Italy — a group renowned for its lack of mobility — to do research in cell biology and related disciplines, taking advantage of the

institute's close links with clinical departments.

Jacopo Meldolesi, the director of research, is determined to break the mould of Italian science by hiring the best people and creating the best possible research environment without feeling an obligation to make appointments on the basis of loyalty to any scientist who has survived long enough in a particular institute. But that system is exactly what geneticist Fulvio Mavilio, who left Naples to join the Dibit Institute, hopes to avoid.

Unhappy with the way research decisions are made in southern Italy, Mavilio is optimistic that Dibit will be different. "The idea here is to focus [on a scientist's] performance, measured in terms of international publications", he says. "Those are criteria not necessarily shared by university teachers or the national research institute here."

Meldolesi, whose own interest in signal transduction and second messenger systems forms one of the major research projects at the new institute, hopes in the next two years to double the current number of scientific staff to around 200. More than a third of the institute's research budget comes from outside sources, including the national research council and the European Communities. Further financial independence is provided



Jacopo Meldolesi

by the drug companies Roche and Bayer, both of which intend to rent space in the institute for fundamental research on drug discovery. Each industrial unit will employ some 30 to 40 scientists.

Italy lags behind such countries as Germany, France and Britain in biology, not to mention the United States. Meldolesi's dream is to create a facility with a reputation to match those of the Institut Pasteur or the Max Planck institutes. **Alison Abbott**

Five-year plan cuts funding for Indian laboratories

New Delhi. India's Council of Scientific and Industrial Research (CSIR) has asked 40 of its laboratories to extend some projects and cancel others after the council's budget for the next five years was cut drastically.

The eighth five-year plan from the Indian Planning Commission approved only £117 million for the council, less than half of its request for £260 million for 1992–97. Moreover, the £9 million allocated for the current year is the lowest in the history of the council, which has enjoyed a steadily rising budget since its formation in 1950. The current budget is enough for only two or three projects in each laboratory.

As a result, many projects beginning this year have been cancelled, while others were either delayed or terminated ahead of schedule. Those abandoned include a high-temperature superconductivity project at the National Physical Laboratory and a very-large baseline interferometry project on studying crustal movements to predict earthquakes.

A £32-million demonstration plant to demonstrate an indigenously developed technology to generate power through coal gasification has been scrapped, as has a project to convert natural gas into diesel oil using a catalyst developed by CSIR scientists.

The government justified its reduced allocation on the grounds that laboratories should find industrial support for some of their programmes; last year that support totalled £75,000. But S.K. Joshi, director general of the council, says that the money usually goes towards specific projects and is not available for the council's own efforts. In addition, the government has weakened the council's ability to earn money through the sale of technologies and from royalties by encouraging industries to import what they lack.

In a separate development, the Indian government has decided to shut down the world's deepest underground physics laboratory, ending almost 30 years of research on cosmic ray neutrinos and the world's

longest-running experiment to detect proton decay.

The government-owned Bharat Gold Mines Ltd decided that further mining of the 112-year-old Kolar gold fields would be uneconomic; annual gold production at the mines has dropped from 40 tonnes two decades ago to 700 kg. Its decision to cease mining by March 1993 presented laboratory officials with an annual bill of £2 million (US\$3.8 million) to maintain and operate the shafts, lifts, air conditioning and ventilation systems. But the Tata Institute of Fundamental Research (TIFR), which operates the laboratory, decided to close the facility after failing to win assurances that the government would provide the additional funding.

The gold mine is 4 km deep and the laboratory extends more than half-way down. In 1965, it was the first to detect interactions of cosmic rays neutrinos, and in 1980, in a collaboration with Japanese scientists, it became the first site to search for proton decay. **K.S. Jayaraman**

Chinese researchers told to fend for themselves

Beijing. China wants two-thirds of the scientists it now funds to find other sources of support as part of the movement towards a free-enterprise system. But many of the 80,000 researchers at hundreds of government-funded institutes believe that the policy fails to account for the varying appeal to industry of some of their work.

They are concerned that the new policy, if not modified, will mean that researchers in some fields will be forced to abandon their profession. While researchers at the Changchun Institute of Applied Chemistry within the Chinese Academy of Sciences, for example, can be expected to fend for themselves, those at the Geosciences Institute and the Desert Research Institute will have a much harder time of it.

The new philosophy was spelled out in late August by the State Science and Technology Commission. Its goal is to give China a modern research and development network that is well-balanced, efficient and vigorous. Its elements include diversifying talent, stabilizing basic research, increasing the number of high-technology projects, creating high-technology enterprises and preserving existing strengths in basic and applied research.

China's policy-makers believe that new

initiatives such as the Climbing programme (see *Nature* 359, 177; 1992) and the 863 high-technology programme are sufficient to keep China at the forefront of global science. Zhou Guang-Zhao, president of the Chinese Academy of Sciences, is fond of saying that these efforts will allow China "to win gold medals in the world science olympics".

But Zhou's optimism is not shared by most working scientists, who can be overheard at conferences talking about opening restaurants and running shops after state funding is cut off. In an effort to remain viable, the National Geology Library of China has rented out half of its building to a local company, for a fee of 2 million yuan (US\$350,000). Unfortunately, the loss of such an important resource can only hurt those left to practise science.

It is not clear whether the new policy is part of a long-term strategy or merely a short-term expedient to withstand the current worldwide recession. Although the government has promised to increase significantly the salaries of those who retain funding, many researchers say that the policy raises questions about the government's commitment to science at any price.

You Qin Li

Japanese increase spending on universities

Tokyo. One per cent of a supplementary budget approved last week by the Japanese cabinet to boost the faltering Japanese economy will be spent on buildings and facilities at government research organizations.

A senior official of Tokyo University, Japan's leading national university, says the supplementary budget is "wonderful news" but that no spending decisions have been made. The universities received a similar windfall in 1987, at the time of the last major supplementary budget, and much of the money was invested in supercomputers and new buildings. The new allocation is for the fiscal year that ends on 31 March 1993.

The biggest beneficiary of science's ¥113 billion (nearly US\$1 billion) share of a one-time ¥10,700 billion supplementary budget is the Ministry of Education, Science and Culture, which will get an extra ¥82.5 billion for the national universities, university-related research institutes, private universities, technical colleges and museums. This is separate from ¥20 billion a year for the next five years already set aside in the ministry's budget to rebuild national universities (*Nature* 355, 99; 1992) with money

derived from selling land owned by them.

The Science and Technology Agency (STA) gets an extra ¥17.8 billion, of which ¥5 billion is for buildings and facilities, including ¥1.5 billion for imported scientific equipment. The remaining money will go to several projects, including the world's most powerful synchrotron, Spring-8, being built in Harima science park, installing seismographs 2,000–3,000 metres underground for earthquake prediction, the Japanese contribution to the US space station and STA's marine and radiological research institutes.

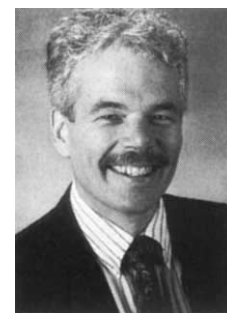
The Agency of Industrial Science and Technology, the main research arm of the Ministry of International Trade and Industry, plans to spend its ¥12.7 billion to shorten by one year construction of two institutes in Tsukuba science city that are part of a reorganization of four of the agency's institutes there (*Nature* 351, 90; 1991). The two institutes — for materials science and life science — had been expected to be completed in fiscal year 1994. Some of the new money will be used for laboratory equipment and cooling systems that do not contain chlorofluorocarbons.

David Swinbanks

NIH plans trials of controversial AIDS drug

Washington. Bowing to pressure from African-American activist groups, the US National Institutes of Health (NIH) last week agreed to go ahead with government-sponsored clinical trials of oral interferon alpha in AIDS patients.

The government has been accused of ignoring or suppressing evidence that the drug, which goes under the names Kemron, Immunex or Immuviron, is effective in treating AIDS symptoms (see *Nature* 359, 661; 1992).



Jack Killen

It has become a cause célèbre for African-American groups, led by the Nation of Islam, after some early uncontrolled clinical trials in Kenya generated reports of dramatic but subsequently un-

firmed results, including seroconversion.

NIH officials met researchers and African-American doctors who have been working with the drug and announced afterwards that the evidence — both scientific and anecdotal — justified taking interferon alpha to full clinical trial. The next step, according to Jack Killen of the National Institute of Allergy and Infectious Diseases, is a plan to conduct the trials, to be developed in the next few months with help from an advisory panel. Current thinking is for a minimum of three sites at a cost of several million dollars. Barbara Justice and Abdul Alim Muhammad, two doctors affiliated with the Nation of Islam who have been prescribing interferon alpha to their AIDS patients for the last year, are expected to be members of the panel.

Despite NIH's approval, the issue remains controversial. Much of last week's closed meeting, according to participants, was focused not on data but on accusations of "genocide" and racism, some levelled at NIH officials.

NIH based its decision to proceed on data provided by Wilbert Jordon of King-Drew Medical Center in Los Angeles. Jordon reported that, in head-to-head trials of four groups taking AZT, low-dose Kemron, a placebo, or a combination of AZT and Kemron, only the Kemron-only group showed improved CD4 counts after one year. Although the data need further examination, they offer NIH enough evidence to justify taking the next step and relieving, for the moment, the political pressure on the agency.

Christopher Anderson

Coping with toxic pulses

SIR — In their recent Commentary article, Tate and Enneking¹ rightly draw attention to the potential hazard of supplying varieties of vetch seeds (*Vicia sativa*) for human consumption as a substitute for red lentils (*Lens culinaris*). Whereas the article focused on the presence in *Vicia* seeds of the toxic amino acid β -cyanoalanine and its γ -glutamyl derivative, there is an even more serious problem associated with the relatively high content of the pyrimidine glucoside vicine in these seeds.

Vicine has been recognized for years as the causative dietary factor in the disease favism, which is characterized by a potentially fatal haemolytic anaemia². Although it is now known to be present in most species of *Vicia*, it was from vetch seeds (*V. sativa*) that vicine was first isolated, a yield of 0.35 per cent being obtained^{3,4}. Favism occurs only in genetically predisposed individuals deficient in erythrocyte glucose-6-phosphate dehydrogenase and, although bouts can be induced by a variety of drugs, they most commonly follow ingestion of broad beans (*V. faba*). In the geographical areas where favism is prevalent, for example the Mediterranean countries, North Africa and the Middle East, sufferers quickly learn to avoid *Vicia* species. Passing off split seeds of *V. sativa* as red lentils in these areas would be particularly unfortunate.

E. G. Brown

Biochemistry Research Group,
University College of Swansea,
Swansea SA2 8PP, UK

SIR — The consumers of the Middle East and India have thousands of years of experience in the preparation of food-stuffs, much of which, if consumed untreated, would be toxic¹. Consumers throughout the Middle East can certainly distinguish between lentils and vetch. The differences in shape (lentil versus spherical), taste (vetch is favoured in some markets for its bean taste), texture and cooking time are natural inbuilt safety factors against substitution.

The latest Australian Bureau of Agricultural and Resource Economics (ABARE) production estimate of *Vicia sativa* for 1992–93 is 75,000 tonnes. It is expected that approximately 60 per cent of this production will be of the cultivar blanch fleur. The rapid expansion of the vetch industry in Australia has meant that the industry has needed to act quickly to address concern about description and quality. This has resulted in the following action by the industry in conjunction with various government departments.

(1) Australian standards for receipt

and export of vetch have been adopted by the industry.

(2) *V. sativa* for export is correctly labelled after phytosanitary testing by the Australian Quarantine Inspection Service, the Australian government body responsible for export inspection services. It has to be labelled as either vetch, blanch fleur vetch or red split vetch.

(3) The Grains Council of Australia (representing all grain growers) has agreed to the introduction of a compulsory research levy on the production of vetch. This will be matched by the Australian government and administered by the Grains Research & Development Corporation (GRDC).

(4) The industry has commissioned research on the levels of neurotoxin in vetch at the Australian Academy of Grain Technology, which ascertained the levels of neurotoxins, and concluded that these levels are reduced by up to 90 per cent through soaking.

(5) The industry has commissioned further research at the academy, under the direction of academy director Dr R. B. H. Wills, on the rate of loss of toxins during soaking and cooking. The nature of breakdown products from the toxins will also be investigated.

The issue of the suitability of blanch fleur for human consumption has been clouded by statements and innuendo from various interest groups on consumption patterns, possible substitution for lentils and the extrapolation of historical data. The industry view is that the opportunity for substitution of vetch (*V. sativa*) for red lentils (*L. culinaris*) is minimal due to the Australian government requirements of appropriate labelling for export and the inherent characteristics of vetch. Differences in taste, texture and cooking time make the two readily discernible to consumers.

The Australian grain legume industry has not been able to identify any published example of human poisoning by *V. sativa*. While we accept there are problems in certain sectors of the livestock industry when excessive amounts of *V. sativa* are included in rations, to extrapolate these findings to humans would endanger the credibility of the proponents of such action.

Allan D. McCallum

Australian Grain Legumes Committee,
Queen Victoria Terrace,
Canberra, ACT 2600, Australia

1. Tate, N. E. & Enneking, T. *Nature* **359**, 357–358 (1992).
2. Mager, J., Chevon, M. & Glaser, G. In *Toxic Constituents of Plant Foodstuffs* (ed. Liener, I. E.) 2nd edn 265–294 (Academic, New York, 1980).
3. Ritthausen, H. & Kreisler, U. *J. prakt. Chem.* **2**, 833–840 (1870).
4. Ritthausen, H. *J. prakt. Chem.* **24**, 202–211 (1881).

Test-ban treaty

SIR — Your leading article (*Nature* **359**, 465; 1992) correctly points out that the conclusion of a comprehensive test-ban treaty would improve the prospects (already quite good) for a successful outcome of the 1995 conference that will decide how long the Nuclear Non-Proliferation Treaty (NPT) will be extended. You make much of the point that France and China should join. In fact China acceded to the NPT in March this year and France followed suit some four months later.

You suggest that a comprehensive test-ban might be brought about by amending the NPT. As a former member of the staff of the IAEA (International Atomic Energy Agency), I should like to point out that it would be far simpler to conclude a separate treaty, replacing the existing partial test-ban treaty of 1963 (which France and China have not joined). Amending the NPT is almost impossible. Every party to the NPT that also happens to be a member of the governing body of the IAEA (there are usually about 28 NPT states on the IAEA governing body, from all parts of the world) can block an amendment by voting against it and again later by failing or forgetting to ratify an approved amendment. No other international arms control treaty places such a formidable barrier in the way of its amendment

and, unsurprisingly, none have been attempted.

David Fischer

15 Willow Walk,
Cambridge CB1 1LA, UK

Thinking physics

SIR — I endorse Sacks' description¹ of Edelman's theory of consciousness as having "great force and originality". Nevertheless, it is not a fully general or "transparent" theory in the sense alluded to by Gray². Edelman³ is concerned primarily with the evolution and neurology of human consciousness. This perspective cannot lead to general predictions about the existence of consciousness in other creatures or in artefacts, as Edelman concedes. The fundamental problem remains: to characterize those physical structures that are necessary and sufficient for the presence of consciousness wherever it might be found. What physics will be required to answer this question remains to be seen.

Ian N. Marshall

57 Bainton Road,
Oxford OX2 7AG, UK

1. Sacks, O. *Nature* **358**, 618 (1992).
2. Gray, J. *Nature* **358**, 277 (1992).
3. Edelman, G. *Bright Air, Brilliant Fire* (Penguin, London, 1992).

Sex education and AIDS

SIR — It was slightly curious to find an editorial attack on the AIDS education policy of the New York City Board of Education made in the name of 'science and medicine' (*Nature* 359, 2; 1992). The board's statement that "abstinence is the most appropriate premarital protection against AIDS" is an incontrovertible medical truth. There is, however, another important fact, as your leading article rightly points out, which is that large numbers of young people engage in premarital sex. Therefore there clearly needs to be effective education also on the prevention of infection during intercourse. But it would not be scientifically correct to pretend that preventive measures are as safe as abstinence, and it would be irresponsible to withhold this important piece of information (which could save someone's life).

The aim of good education is to give people the facts and then let them make up their own minds. There is a certain irony in a scientific journal campaigning to prevent people being told clearly established facts.

Denis R. Alexander
Institute of Animal Physiology
& Genetics Research,
Cambridge Research Station,
Babraham,
Cambridge CB2 4AT, UK

SIR — You are unfair to the New York City Board of Education. There is no scientific study proving that an educational programme emphasizing the use of condoms will protect a specified population against AIDS. In fact, in the school-child population being targeted in New York City, much medical common sense suggests that emphasis on the use of condoms will cause more AIDS than it will prevent.

First, as contraceptive experience has shown, condoms have at least a 2 per cent failure rate (as barrier to semen), because of breakage and slippage. Second, as students of teenage sexuality have shown, consistent use of condoms with consecutive coituses is extremely low in the teenage population. Third, emphasis on the use of condoms alone ignores the fact that AIDS is a disease spread by all body secretions and excretions, and not just by semen. Condoms do not protect against AIDS risks in open mouth-to-mouth kissing, body-biting and much oral sex play.

On top of all these potential failures of the condom to protect against an infection with virtually 100 per cent mortality, the implicit acceptance by an educational institution of the legitimacy of its young teenage students having 'safe sex' in a society already made hypersexual by

media stimulation can serve only to increase all kinds of sexual contact among the teenage population and therefore increase the spread of AIDS.

Saying no to too early sexual activity ought to be the chief topic of education against AIDS in such young school populations. It is the only 'safe sex'. The New York City Board of Education should be congratulated for showing such intelligent caring for its young wards. One would wish that *Nature* and New York's Mayor Dinkins would be as caring.

Edward Stim
Takashimadai, 4-4,
Nagamine Building, Apt 302,
Yokohama 221, Japan

Anal sex and AIDS

SIR — Per-Erik Åsard¹ exhorts the AIDS scientific community to broadcast the dangers of anal sex, as practised by heterosexuals and bisexuals. It has been shown² that the relative risk of human immunodeficiency virus (HIV) transmission from men to women is increased by a factor of 5.1 in anal compared with vaginal intercourse: that is, 46 per cent of women who practised anal sex with an HIV-infected man became HIV-positive.

Such appeals might be more effective if reinforced with mechanical rather than moral or statistical arguments. Here are some: the colonic and rectal mucosa has a barrier function that normally prevents overwhelming invasion by infective and toxic materials contained within the luminal contents. Exposure of rat colonic mucosa for 2–3 hours to human semen causes extensive damage to the rat mucosa as shown by histology, inhibition of fluid absorption and an increase in permeability to the normally impermeant high molecular weight probe, polyethylene glycol 4000 (ref. 3).

Human semen contains at least two components in sufficiently high concentrations to cause breakdown of the basement membrane that supports the colonic epithelial cell layer: collagenase, a component of seminal plasma, which hydrolyses the collagen present within the basement membrane, and spermine (concentration range in human semen 10–15 mM). Spermine permeates through the colonic mucosa and neutralizes glycosaminoglycans within the interstitial matrix — this triggers the activation of endogenous collagenases which leads to loss of mucosal barrier function — this allows seminal collagenase to penetrate the mucosa and hence to cause further damage and breakdown of the barrier function. Most of the mucosal

damage caused by semen can be prevented by synthetic inhibitors of collagenase⁴. Thus it is apparent that the colorectal mucosa is particularly susceptible to biochemical as well as the normally assumed mechanical trauma consequent upon anal intercourse⁵.

Richard J. Naftalin
Biomedical Sciences Division,
Physiology Group,
King's College London,
Strand, London WC2R 2LS, UK

1. Åsard, P.-E. *Nature* 359, 182 (1992).
2. European Study Group on Heterosexual Transmission of HIV. *Br. med. J.* 304, 811 (1992).
3. Mendizabal, M. V. & Naftalin, R. J. *Clin. Sci.* 82, 277–282 (1992).
4. *Clin. Sci.* 83, 417–423 (1992).
5. Lorian, V. *Lancet* i, 1111 (1988).

Belgian science

SIR — As Alison Abbott says (*Nature* 358, 445; 1992), Belgium is divided both into two cultural communities, French-speaking and Flemish-speaking, and into three regions Flanders, Wallonia and the capital Brussels. Each of these divisions serves a different administrative purpose, and basic science is administered at the community level. Thus the figure comparing the funds allocated by the Flemish and French-speaking divisions of the National Fund for Scientific Research (NFSR) to the populations of Flanders and Wallonia is misleading because it neglects the population of the Brussels region; the figures should instead have been compared with the French- and Flemish-speaking populations, including the one million citizens of Brussels. This is very important because 80 per cent of the capital's population is French-speaking.

Second, it was stated that a 'French speaker cannot apply for a grant from the Flemish NFSR nor vice versa'. In fact, the sole criterion determining which section of the NFSR is relevant is the language used by the university at which the applicant studied. At the moment, any Belgian is free to choose any university, irrespective of language, though the Flemish government is now trying to change this.

May I also point out that the Belgian state, which has its roots in the Middle Ages, was created in 1830, not the "1830s". It seceded from the Netherlands after the Belgian revolution. It is therefore tendentious to qualify Belgium as "the product of recent history" obtained by "throwing together" Flanders and Wallonia and to limit its *raison d'être* to a "buffer between France and Germany".

G. Berben
Station de Chimie
et de Physique Agricoles,
B-5030 Gembloux, Belgium

The infectiousness of pompous prose

Martin W. Gregory

For centuries, scientists have been bombarded with pleas for plain language. Why have these pleas had no effect, when the problem of unreadable prose could be solved at a stroke?

THERE are two kinds of scientific writing: that which is intended to be read, and that which is intended merely to be cited. The latter tends to be infected by an overblown and pompous style. The disease is ubiquitous, but often undiagnosed, with the result that infection spreads to writing of the first type. I would like to present a few examples of the problem, and to offer a solution.

In 1667, Bishop Thomas Sprat implored¹ the newly formed Royal Society of London for the Improving of Natural Knowledge to "reject all the amplifications, digressions, and swellings of style; to return back to the primitive purity, and shortness, when men delivered so many *things*, almost in as many *words*." I have quoted only one sentence, but the bishop, evidently used to a captive audience, took two pages to extol brevity.

The most common problems that occur in scientific writing are (1) too many words, and (2) the adoption of a supposed 'literary' style in the mistaken belief that the written language is different from the spoken.

A typical example of the first problem is shown in the figure. The following is another: "The *main* purpose of any scientific article is to convey in the fewest *number* of words the ideas, procedures and conclusions of an investigator to the *scientific community*. Whether or not this *admirable* aim is accomplished depends to a *large extent* on how skillful the author is in *assembling the words* of the *English language*."

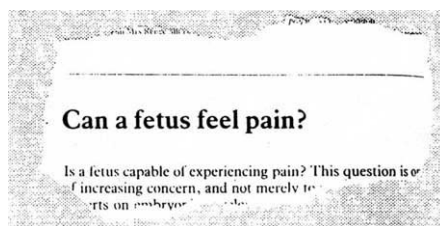
That is the opening sentence of an editorial² entitled "Use, misuse and abuse of language in scientific writing". The italics are mine: all these words can be omitted without loss of meaning. The last sentence can be reduced from twenty-seven words to eight: "Whether he succeeds depends on his writing skill." Thus the problem is so insidious that it appears even in works devoted to its eradication.

Distorted prose

In the second type of problem, prose is perverted and distorted to make it difficult to understand. It is like a neoplastic transformation, rendering the original tissue (the spoken word) unidentifiable. To illustrate this point, I shall present a case, and ask you to imagine yourself using it to explain your work to a

colleague in the bar. Try opening the conversation with: "The availability of culture methods to measure either the formation of antibody in mixtures of T and B cells or the antigen-driven proliferation of T cells has allowed a more precise evaluation of the phenomenon involved in induction of the immune response³." If I have understood the author correctly, he means: "Some culture methods have helped us to understand the immune response better. Using cell culture we can measure antibody formation in a mixture of T and B cells — or we can measure the proliferation of T cells driven by antigen."

This may not be much shorter, but it is



A stylistic problem⁴. The first sentence merely repeats the title, but instead of being omitted it has been doubled in length.

more likely to get you a drink. Try this one: "That the sense of smell was used by these cattle was established because of the marked audible variation in inhalation intensity as the animals grazed⁴." Presumably, "marked audible variation in inhalation intensity" means loud sniffing.

Here is a severe case: "As the practical relevance of intestinal immunity in diarrhoeal disease relates to the possibility of developing effective immunisation programmes for the control of gut infections, this review will focus on insights into the functioning of the immune system particularly relevant to this goal⁵." In other words: "This review will focus on aspects relevant to vaccine development."

Aaronson wrote an article⁶ called *On Style in Scientific Writing*. He cited C. D. Graham, who compiled a glossary of pompous phrases, of which Aaronson gave some examples. Rather than cite any of these examples, I quote Aaronson's comment on them: "Although Graham is pressing the point for the sake of humor, working scientists will recognise the essential veracity of his translations." Had Aaronson been talking, he would have said "he's funny but he's right".

If you wish to be unintelligible, start your sentences in the middle so that the reader doesn't know what you're talking about until half-way through: "Similar to figurative language in function, humor is another way by which we come to know the world." That quotation was taken from the book *Breathing Life into Medical Writing*⁷, in which this abominable style is actively encouraged. With skill, the technique can be refined to the point where the reader has to go back to the beginning to find out how it all started: "No avicide myself and, indeed, not much of a wide-ranging aviphage, I had always assumed that the so-called glorious twelfth occurred only in August when the aristocratic victim of your matched pair of Churchills or Boss' or Purdies (or what have you at £12,000 a throw) is, of course, our only and uniquely indigenous bird, the red grouse⁸."

That sentence deserves a prize. Read enough times, it becomes apparent that the author is not talking about avicides, or about himself, or about wide-ranging aviphages, or about aristocratic victims or about throwing matched pairs of Churchills, Bosses or Purdies (whatever they are), or even about the red grouse. He's talking about *the so-called glorious twelfth*. What would a non-British reader make of all that?

When a paper of this type is read aloud, the effect is stunning. In 1880, T. H. Huxley is said to have opened his speech to the Zoological Society with this sentence: "There is evidence, the value of which has not been disputed, and which, in my judgement, amounts to proof, that between the commencement of the tertiary epoch and the present time the group of the equidae has been represented by a series of forms, of which the oldest is that which departs least from the general type of structure of the higher mammalia, while the latest is that which most widely differs from that type⁹."

The titles of some scientific journals are admirably brief. Not so the contents. Here is an example from *Gut*: "All of these measurements have wide ranges of values in both control (Doniach and Shiner, 1957; Butterworth and Perez-Santiago, 1958; Rubin *et al.*, 1960a and b; Shiner and Doniach, 1960; Chacko, Job, Johnson, and Baker, 1961; Cameron *et al.*, 1962; Jos, 1963; Yardley, Bayless, Norton, and Hendrix, 1962;

Astaldi, Conrad, Ratto, and Costa, 1965; Madanagopalan *et al.*, 1965; Swanson and Thomassen, 1965; Stewart, Pollock, Hoffbrand, Mollin, and Booth, 1967; Pollock, Nagle, Jeejeebhoy, and Coghill, 1970) and coeliac (Rubin *et al.*, 1960a; Shiner and Doniach, 1960; Chacko *et al.*, 1961; Cameron *et al.*, 1962; Jos, 1962; Yardley *et al.*, 1962; Bolt, Parrish, French, and Pollard, 1964; Madanagopalan *et al.*, 1965; Stewart *et al.*, 1967; Hamilton, Lynch, and Reilly, 1969; Pollock *et al.*, 1970) mucosae and the differences between the means are small (Rubin *et al.*, 1960a; Shiner and Doniach, 1960; Jos, 1962; Madanagopalan *et al.*, 1965; Stewart *et al.*, 1967)¹⁰."

That is not verbosity. It is simply the unspeakable Harvard system, to which so many journals are needlessly committed. The 149 words can be reduced to 22 at a stroke: "All these measurements have wide ranges of values in both control¹⁻¹² and coeliac^{3-7,9,11-14} mucosae and the differences between the means are small^{3,4,6,9,13-15}". The original passage is unspeakable and unreadable, but neither the author nor the editor is interested in whether anyone *reads* this article. Indeed, they prefer that no one reads beyond the summary, or better still, beyond the authors' names.

Treatment

The first of these stylistic problems is easy to treat. Authors need to reduce their articles to the fewest possible words. William Strunk dealt with the subject well in his classic *The Elements of Style*¹¹. Strunk was an enthusiast. One of his students described his zeal in a later edition¹² of Strunk's book: "Omit needless words!" cries the author on page 23, and into that imperative Will Strunk really put his heart and soul. In the days when I was sitting in his class, he omitted so many needless words, and omitted them so forcibly and with such eagerness . . . that he often seemed in the position of having shortchanged himself — a man left with nothing more to say yet with time to fill . . . Will Strunk got out of this predicament by a simple trick: he uttered every sentence three times. When he delivered his oration on brevity to the class, he leaned forward over the desk, grasped his coat lapels in his hands, and, in a husky, conspiratorial voice said, 'Rule Seventeen. Omit needless words! Omit needless words! Omit needless words!'" This advice is easy to follow: all you need is a blue pencil and practice.

Treatment of the pseudo-literary style problem is difficult, and rarely attempted, even though editorial boards of journals could solve it at a stroke by rejecting incomprehensible manuscripts. The best advice to authors is to throw the draft away and start again. How can

a plain, clear text emerge by this process? The scientific literature itself will provide little guidance. Some of it may be good, but in my view the best advice is to be found elsewhere. John Whale, for instance, wrote a series of articles for the *Sunday Times* which has been gathered into one volume¹³ which I found particularly readable. In chapters 5 and 6 he urges us to write as we would speak. If this rule is followed, the problem virtually disappears. The solution



T. H. Huxley — overlong declamations?

may be simple, but that doesn't mean it is always easy to apply. You will no doubt find signs of the problem in this paper: I am only a scientist, after all.

Discussion

Pleas for scientists to write readably have failed for at least 300 years. This is because the pleas have been aimed at the wrong people: the scientists. They should have been aimed at editors. The most important aspect of a scientific paper is its scientific value, but no matter how important it is, no one will read it if it is unreadable. Most scientists have no expertise in writing. We need help. The people who should be best qualified to help us are the editors through whom our efforts pass on the way to publication. But whom do we find as editors? More scientists!

Everyone can write, so it is assumed that writing is easy, or unimportant. Everyone can paint as well, but not everyone's paintings are worth hanging on walls. To expect scientists to produce readable work without any training, and without any reward for success or retribution for failure, is like expecting us to play violins without teachers or to observe speed limits without policemen.

Some may do it, but most won't or can't.

With no guidance, scientists copy what they see, and we see things like this: "The author is of the opinion that it is appropriate to write scientific papers in the third person." This is ridiculous. *I* am the author, not a third person.

If there must be scientists on editorial boards of journals, their job must be only to reject bad science. The editor's job is to reject bad writing. The editor has the power to enforce standards of readability, and should be allowed to use this power.

In conclusion, my suggestion for the elimination of unreadable papers is first to omit needless words, and second, to write as you would speak. This magnificent advice will certainly have no measurable effect. For centuries, we scientists have been showered with advice on how to write readably, and still we all ignore it. Isn't it time we sought another solution?

In my opinion, editors should be writers, not scientists. Scientists should be judged according to how many times their work is *read*, not cited. (Audience research reveals how many people listen to or watch which programmes; readership research could in principle reveal how many read our articles.) Peer review will continue to uphold scientific standards, but badly written work should be rejected, whatever its scientific standard. Authors will be prepared to work (or pay) to get their paper re-written to have it published in the best journals if the only grounds for rejection by the journals are those of unreadability. An editor could offer to rewrite such articles (for an exorbitant fee). Thus, editors will get rich, journals will get read, readers will retain their hair, and real progress will be made. □

Martin W. Gregory is in the British Veterinary Project, Republic of Yemen, and can be contacted at the Foreign and Commonwealth Office, King Charles St, London SW1A 2AH, UK.

1. Sprat, T. from *The History of the Royal Society of London for the Improving of Natural Knowledge* (1667) reprinted in Ryan, L. V. *A Science Reader* (Rinehart, New York, 1959).
2. Gaafar, S. M. *Vet. Parasit.* **8**, 199–200 (1981).
3. Unanue, E. R. *New Engl. J. Med.* **303**, 977–985 (1980).
4. Arnold, G. W. & Dudzinski, M. L. *Ethology of Free-Ranging Domestic Animals* (Elsevier, Oxford, 1978).
5. Rowley, D. & La Brody, J. *J. Diarrhoeal. dis. Res.* **4**, 1–9 (1986).
6. Aaronson, S. *Curr. Cont. Life Sci.* No 2, 10 (1977).
7. Sheen, A. P. *Breathing Life into Medical Writing* (Mosby, St Louis, 1987).
8. Hillaby, J. *New Scientist*. (6 May 1976).
9. Huxley, T. H. *Science* **2** (1st series, 1881). (Cited in *Science* **207**, 750: 1980).
10. Chapman, B. L., Henry, K., Paice, F., Stewart, J. S. & Coghill, N. F. *Gut* **6**, 905–909 (1973).
11. Strunk, W. *The Elements of Style* (1919). (Reprinted by Macmillan, New York, 1959.)
12. White, E. B. in *The Elements of Style* (ed. Strunk, W.) xiii (Macmillan, New York, 1959).
13. Whale, J. *Put it in Writing* (Dent, London, 1984).
14. Richards, T. *Br. Med. J.* **291**, 1220–1221 (1985).

Disease-of-the-month alive and well

Identifying diseases in need of a cure is easy and the temptation to follow that with money appears to be irresistible, but it has not always proved to be successful. Research money should follow scientific opportunity.

In the United States and several European nations, this is the "Decade of the brain". In the United States it is also a decade recently dedicated to "Women's health". Since the National Cancer Act of 1971, when the White House declared a full-blown "war on cancer", the US National Institutes of Health (NIH) has been gamely searching for strategies to stop the unyielding progression of tumour cells from disease to death but all too often the tumours are still winning. So is the phenomenon we call ageing, despite the creation in 1976 of a National Institute on Aging whose scientists are tackling everything from the normal deficiencies that decades of wear and tear inflict on the body to the mental ravages of senility and Alzheimer's disease.

The desire to cure disease (and to prevent it in the first place) may be one of human beings' most basic instincts. It is laudable and not without success. Particularly in the developed nations of the world, where sanitation, vaccination and antibiotics, as well as modern surgery, are generally available, numerous diseases can be brought under control. And even for the more recalcitrant diseases, modern medicine can offer a greatly enhanced quality of life for longer periods of time than ever before. But the scientific search for understanding is far from complete (perhaps adolescent, at best).

In this environment, the increasingly crass politicization of biomedical research looms as something of a menace for it presupposes first that more money for more scientists is the way to medical salvation, and second that well-meaning groups of citizen-activists and professional lobbyists have a scientifically useful role in deciding where research money should be directed. There is, alas, inadequate evidence to support either proposition. Recent history illustrates a point that should have been learned long ago.

■ **Breast cancer.** An influential group of activists organised as the Breast Cancer Coalition lobbied so successfully that Congress has allocated \$210 million of Department of Defense money for breast cancer research (see *Nature* **359**, 764; 1992). Coming in addition to \$197 million for breast cancer already in the budget of the National Cancer Institute, the \$407 million total leaves basic researchers in the unenviable position of actually having more resources than can necessarily be well spent on basic science while other programs go

wanting (see page 2). If the grand effort does not yield striking results, there may be political hell to pay. (Even the most zealous activists prepared a research agenda that would spend "only" \$300 million for breast cancer studies.)

■ **Breast cancer and women's health.** As this journal noted recently (see *Nature* **359**, 760; 1992), under the mantle of its "Women's health initiative," the NIH plans to spend \$10 million on a large epidemiological study of the putative causal relation between breast cancer and dietary fat despite several good studies that have disappointingly shown no connection. But some women's groups appear to be unconvinced, and the faddish belief that all disease is directly related to diet seems to have overcome not only scientific data but common sense as well.

■ **AIDS vaccine.** In one of the more egregious examples of doing science by lobbying, a small biotechnology firm called MicroGeneSys in Meriden, Connecticut, managed to get Congress to pass legislation mandating that the Department of Defense allocate \$20 million for clinical trials of its AIDS vaccine (based on the gp 160 envelope of the human immunodeficiency virus), despite an absence of scientific consensus that this particular experimental vaccine is more promising than others. Bernadine Healy, director of the NIH, along with Food and Drug Administrator David Kessler, are now in the politically awkward position of trying to stop this lobbyist-driven clinical trial from proceeding. (They are right to do so.)

■ **AIDS drug.** Valuable research dollars are about to be spent by NIH on studies of an interferon-based drug called Kemron that is popularly known as the "African AIDS cure" because of its association with studies in Kenya. Despite the fact that an NIH review panel has already investigated and rejected data purporting to show that Kemron is a significant AIDS remedy, racial politics demand yet another look and NIH likely will be hard pressed to resist (see *Nature* **359**, 660; 1992).

These are but a handful of instances that demonstrate the power of politics over biomedical science. The history of NIH shows that, to some extent, it was ever thus. After the White House declared war on cancer in 1971, it elevated heart disease to special status (and funding) a year later. During the 1970s the prevailing disease-of-the-month mentality that had every disease group touting its

own cause led to the renaming of many of the NIH institutes, rendering their new identities not only unpronounceable but hard to remember as well. The National Heart Institute became the National Heart, Lung and Blood Institute. The arthritis lobby succeeded in getting Congress to establish a separate National Institute of Arthritis, but political pressures stretched it ever further to the National Institute of Arthritis and Musculoskeletal and Skin Diseases.

The bureaucracy has grown (it is estimated that it costs at least \$5 million a year in administrative costs alone to create new institutes) but, although some progress has been made, it is not yet possible to say that any of the diseases subject to special political attention has been cured as a result. The reason for this sad state is intellectually apparent even if politically unpopular, for the truth is that good science depends on good ideas (not good intentions) if it is to be successful. But even then the system does not always work ideally, as the "Decade of the brain" attests. If the impetus behind more research on breast cancer is a desire to stem the apparent rise in the disease, the push for more research in neuroscience is more clearly grounded in scientific opportunity.

During the past decade, important insights in the molecular analysis of cell lineage and neuronal migration make possible new studies of the central nervous system. The way the brain organizes the chemistry of the five senses is close to being understood. And some diseases may even be experimentally treatable with new techniques of gene therapy. But the "Decade of the brain" is little more than an empty declaration because it was not accompanied by the sums of money that Congress has earmarked elsewhere on a disease-by-disease basis.

It all comes down to the fact that in the United States, at least, the system for allocating research resources is badly in need of a complete overhaul. Even a doubling of the budget would not necessarily lead to money well spent unless there is a commitment to assigning research funds according to scientific opportunity. Although the NIH's much maligned "strategic plan" is meant to be a step in the direction of bringing scientific sense into a system run amok, it clearly does not go far enough when it comes to making tough but appropriate decisions.

Barbara J. Culliton

Potential awareness of plants

Keith Roberts

A GANG of caterpillars grazing on a tomato leaf does not go unnoticed. The leaf's unwounded neighbours are rapidly alerted to the insult and, in response, adopt a series of defensive postures that include the synthesis of proteinase inhibitors designed to give the caterpillars, should they move on to pastures new, a bout of indigestion. But how do cells in one leaf send rapid and appropriate signals to those in another leaf on the same plant? For many years the search has been on for the hypothetical agent, produced in the wounded leaf, that might move systemically through the plant, probably in the phloem, to alert distant leaves. Now, work presented by Wildon *et al.* on page 62 of this issue¹ shows that a propagated electrical signal, rather than a mobile chemical, may be the messenger that carries the bad news from leaf to leaf.

But first, the play so far. More than 20 years ago Ryan's group² characterized two proteinase-inhibitor genes (*pin 1* and *pin 2*) and demonstrated that their expression in unwounded leaves was systemically induced by wounding another leaf elsewhere on the plant. Ryan suggested that a chemical factor, called PIIF (proteinase inhibitor-inducing factor), moved through the plant from the wound site to distant leaves³. Since then, Ryan's group and others have introduced us to a series of actors who have all had, at one time or another, ambitions to play the lead role of PIIF, the systemic messenger. Ryan has recently reviewed the full cast and their histories⁴.

Early on it was found that the act of cutting off a stem or leaf cleanly with a razor blade was not in itself enough to elicit a wound response. This suggested a simple assay for compounds that, when applied to the cut stem or petiole, would enter the transpiration stream and trigger the synthesis of *pin 1* and *pin 2* proteins in leaves⁴. Thus, pectic fragments, in particular small oligogalacturonides, entered stage right and topped the bill for a while, until doubt was cast on their ability to move through the plant *in vivo*⁵. Nevertheless, applied exogenously, they do have a potent induction effect that can be reversed by agents, such as aspirin, that affect membrane ion transport⁶, hinting at mechanisms to come.

More recently, three aspiring new players each made their entrance. The first is jasmonic acid, or its volatile ester methyl jasmonate⁷⁻⁹. When applied to leaves, either directly or indirectly through cut stems, these dramatically stimulate *pin* expression at the transcriptional level in a wide range of plants.

They also induce the accumulation of other wound and abscisic acid-responsive transcripts¹⁰, but the mode of synthesis of jasmonic acid suggests it is involved in intracellular regulation and is not the systemic signal. The second player is abscisic acid itself. This plant growth regulator, when applied exogenously, can again induce the characteristic pattern of wound-response genes, including

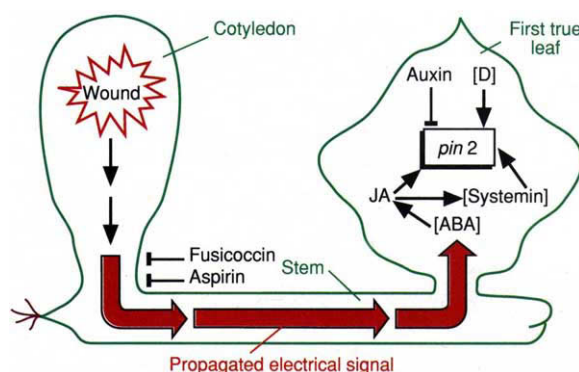


Diagram of how propagated electrical signals might, as part of a signal-transduction pathway, convey information from a wounded cotyledon of a young tomato plant into the first fully-expanded leaf. The order of events within the distant leaf is hypothetical but fits with much of the available data. Steps in brackets may not be common to all plants. Auxin is shown as a repressor of *pin 2* induction, D represents developmental control points on *pin 2* expression, and various inhibitors are shown that act to block the systemic signal. For simplicity, no steps to *pin 2* product are shown in the coleoptile, although it is possible that many of the same steps operate here as in the wounded leaf. ABA, abscisic acid; JA, jasmonic acid.

pin 2 (refs 10,11). However, that very elegant and persuasive set of experiments, which also suggested that methyl jasmonate acts downstream of abscisic acid, was carried out on potato plants, which seem to respond differently to the tomato. The third player is a small polypeptide, 18 amino acids long, called systemin^{12,13}. This molecule elicits features of the systemic wound response at remarkably low concentrations, but it has no signal peptide and so again seems more likely to have an intracellular regulatory role rather than being PIIF itself. So, the stage is set.

Now plant scientists, with a few exceptions¹⁴⁻¹⁶, have been curiously reluctant to entertain the notion that plants, with their symplastic continuity through plasmodesmata, are well suited for electrical signalling. For a variety of reasons, one of which has been lack of reproducibility, action potentials in plants have had a rather bad press. The results presented by Wildon and collaborators in this issue go a long way to

redress this situation, and they make a new and resilient case for awarding the role of PIIF, not to a mobile chemical at all, but to propagated electrical signals that, in many respects, resemble epithelial conduction in lower animals.

Using very young tomato plants, Wildon *et al.* show that wounding the cotyledon (seed leaf) mechanically results in the slow transmission of an action potential out of the cotyledon and into the first leaf, where this correlates in all cases with the induction of *pin 2*. The latter is measured both by a proteinase inhibitor-activity assay, and by the appearance of *pin 2* transcripts. Excision of the unwounded cotyledon gives neither an electrical signal nor *pin* induction, allowing three important experiments to be done.

By removing the wounded cotyledon both just before the electrical signal exits at the petiole and after it has passed, Wildon *et al.* show that without the electrical signal there is no *pin* induction and that with it no further signal needs to pass from the wounded leaf. This still leaves open the possibility of a factor moving rapidly in the phloem (the transpiration stream was excluded) at the same rate as the action potential (1–4 mm s⁻¹). However, by chilling the petiole, it was found that all phloem transport could be blocked, without affecting either the propagation

of the action potential or the induction of *pin*. The crucial experiment was thus to pre-cool the petiole for 5 minutes, wound the cotyledon, wait for the action potential to pass out of the petiole, and then excise the cotyledon. Under these conditions it is very hard to envisage a chemical agent on the move, and yet reproducible induction of *pin* was observed. In a similar way, the authors rule out any role for the rapidly propagated (100 mm s⁻¹) hydraulic 'signals' induced in plants by wounding¹⁷. This effectively leaves the stage clear for the propagated electrical signal to play PIIF, at least in small tomato plants.

Nonetheless a lot of problems remain, and timing is one of them. The time taken for the signal to leave a more mature tomato leaf is reported to be much longer, suggesting that the developmental age of the plant and the leaf that is wounded may affect events considerably. In addition, even in the small young plants, the time taken for induction of *pin* protein is long (2–4 hours)

relative to the passage of the electrical signal, and the sequence of events in the unwounded leaf will now need to be addressed in earnest. Many of the regulatory factors involved have already been mentioned but their ordering is deeply problematic^{18,19}, as is the question of whether a similar or different set of signal transduction steps goes on in the wounded, and the unwounded, leaf (see figure).

Two further complexities also need to be taken into account. The first is the plant growth factor auxin, which has

been proposed²⁰ as an endogenous natural repressor of *pin* expression¹¹. The second is that *pin 2* protein is developmentally regulated as well as being wound-inducible; it is abundant, for example, in floral buds⁸. However things fall into place eventually, the results of Wildon *et al.* will stimulate a flurry of research and it will be some time before the curtain falls on this particular show. □

Keith Roberts is in the Department of Cell Biology, John Innes Institute, Colney Lane, Norwich NR4 7UH, UK.

1. Wildon, D. C. *et al.* *Nature* **360**, 62–65 (1992).
2. Green, T. R. & Ryan, C. A. *Science* **175**, 776 (1972).
3. Ryan, C. A. *Pl. Physiol.* **54**, 328–332 (1971).
4. Ryan, C. A. *Pl. molec. Biol.* **19**, 123–133 (1992).
5. Baydoun, E. H. & Fry, S. C. *Planta* **165**, 269 (1985).
6. Doherty, H. M. & Bowles, D. J. *Physiol. molec. Pl. Path.* **33**, 377–384 (1988).
7. Farmer, E. E. & Ryan, C. A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 7713–7716 (1990).
8. Farmer, E. E., Johnson, R. R. & Ryan, C. A. *Pl. Physiol.* **98**, 995–1002 (1992).
9. Creelman, R. A., Tierney, M. L. & Mullet, J. E. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4938–4941 (1992).
10. Hildmann, T. H. *et al.* *Pl. Cell* **4**, 1157–1170 (1992).
11. Peña-Cortés, H. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 9851–9855 (1989).
12. Pearce, G., Strydom, D., Johnson, S. & Ryan, C. A. *Science* **253**, 895–897 (1991).
13. McGurl, B., Pearce, G., Orozco-Cardenas, M. & Ryan, C. A. *Science* **255**, 1570–1573 (1992).
14. Van Sambeek, J. W. & Pickard, B. G. *Can. J. Bot.* **54**, 2642–2650 (1976).
15. Roblin, G. *Pl. Cell Physiol.* **26**, 455–467 (1985).
16. Wildon, C., Doherty, H. M., Eagles, G., Bowles, D. J. & Thain, J. F. *Ann. Bot.* **64**, 691–695 (1989).
17. Malone, M. *Planta* **187**, 505–510 (1992).
18. Bowles, D. *Curr. Biol.* **1**, 165–167 (1991).
19. Enyedi, A. J., Yalpani, N., Silverman, P. & Raskin, I. *Cell* **70**, 879–886 (1992).
20. Kernan, A. & Thornburg, R. *Pl. Physiol.* **91**, 73 (1989).

ASTRONOMY

Stars at the quantum limit

Paul C. Joss

QUANTUM effects are familiar in the realm of the microscopic, and they are fundamental to the behaviour of semiconductors, but on page 48 of this issue¹ Chabrier, Ashcroft and DeWitt show that they can affect the properties of matters even on the giant scale of stars.

As stars with masses less than about three times the mass of the Sun exhaust the last of their nuclear fuel, they shed their outer layers and shrink down to become white dwarfs. These final cinders of stellar evolution still have masses close to that of the Sun, but only a millionth the volume; their sizes are comparable to that of the Earth, and their central densities can reach 10^9 g cm⁻³ (a billion times that of water). White dwarfs are responsible for all sorts of astronomical phenomena, including cataclysmic variables, novae and type I supernovae, and their luminosity can be used to calibrate the ages of the stellar systems in which they reside.

Because nuclear burning in their interiors has been extinguished, the internal structure of white dwarfs is simpler than that of stars in earlier evolutionary phases. Since the pioneering work of Chandrasekhar², it has been understood that it is primarily the pressure of 'degenerate' electrons (electrons in the lowest-energy configuration available) that supports a white dwarf against collapse from its considerable self-gravity — Heisenberg's uncertainty prin-

ciple prevents the electrons getting any closer together. For more than 30 years, it has also been recognized³ that as a white dwarf radiates and cools, the electrostatic interactions among the fully ionized nuclei in its interior eventually cause the nuclei to freeze into a lattice. The crystallization wave that propagates outwards from the stellar centre has important consequences for the thermodynamics and heat-transport properties of the white-dwarf interior. It is in this freezing phenomenon that Chabrier *et al.* now find significant quantum effects which were previously neglected.

Previous studies of the crystallization of white-dwarf interiors have treated the nuclei in a purely classical manner, in both the fluid and solid phases. This seemed to be a reasonable approximation, because under white-dwarf interior conditions the de Broglie wavelength, λ , of the nuclei is smaller than the mean internuclear separation, r , suggesting that the nuclei can be treated as an ensemble of classical point charges. However, Chabrier *et al.* note that the criterion $\lambda/r < 1$ is not a sufficient condition for quantum effects to be negligible. They show that the irremovable quantum-mechanical zero-point energy per nucleus in the solid phase actually exceeds the thermal energy, so that the nuclei must be treated as a quantum solid. The authors further argue that quantum corrections will be comparably

important in the fluid phase as the freezing point is approached; hence the fluid, too, is a partially quantum liquid rather than a classical one. In both states, the quantum effects raise the free energy by 10–25 per cent (depending on the density and chemical composition) over the value obtained in the classical limit.

Applying a semi-empirical analysis, Chabrier *et al.* estimate that quantum effects lower the melting temperature by about 10 per cent in the most massive white dwarfs, which have masses about 1.4 times that of the Sun. Consequently, these stars must cool considerably more than had been realized before crystallization effects start to set in. Such massive examples are known only as members of close binary stellar systems. The more common isolated white dwarfs typically have much lower masses, clustered near 0.6 solar masses. For these, the estimates by Chabrier *et al.* suggest that the melting temperature is almost unaffected by quantum effects — the quantum corrections are still appreciable but are nearly equal in the solid and fluid phases, and so tend to cancel out.

Despite this cancellation, the quantum corrections may have an important bearing on the cooling history of all white dwarfs regardless of their masses. Because quantum effects are already significant in the fluid phase, Debye cooling (a quantum mechanical phenomenon that lowers the heat capacity of the nuclei) should commence even before crystallization sets in. As a result, a white dwarf will cool more rapidly than had previously been expected. This accelerated cooling will, in turn, decrease estimates^{4,5} of the age of the galactic disk derived from the distribution of luminosities among white dwarfs in the solar neighbourhood, perhaps by $1\text{--}2 \times 10^8$ years.

Although Chabrier *et al.* have convincingly demonstrated the importance of quantum effects on the thermodynamics of the matter in white dwarfs, their estimates are too crude to permit a definitive evaluation of the astrophysical consequences. As they point out, there is now a pressing need for more accurate theoretical determinations of the impact of these quantum phenomena on crystallization and cooling under the conditions prevalent in white-dwarf interiors. □

Paul C. Joss is in the Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.

1. Chabrier, G., Ashcroft, N. W. & DeWitt, H. E. *Nature* **360**, 48–50 (1992).
2. Chandrasekhar, S. *Stellar Structure* (University of Chicago Press, 1938).
3. Salpeter, E. E. *Astrophys. J.* **134**, 699 (1961).
4. Winget, D. E. *et al.* *Astrophys. J.* **315**, L77–L82 (1987).
5. García-Berro, E., Hernanz, M., Mochkovitch, R. & Isern, J. *Astr. Astrophys.* **193**, 141–147 (1988).

Riding high on the TATA box

Jack Greenblatt

AN eagerly anticipated development, the three-dimensional structure of the TATA-box binding protein (TBP) appears on page 40 of this issue¹. TBP is the DNA-binding subunit of the general factor TFIID required for transcriptional initiation by RNA polymerase II. Nikolov *et al.*¹ have used X-ray diffraction analysis to determine the structure at 2.6 Å resolution of one of two nearly identical species of TBP found in the plant *Arabidopsis thaliana*. The results are of universal value because 70 per cent of the 180 carboxy-terminal amino acids of TBP are identical in all eukaryotes studied so far (including plants, mammals, insects and fungi); also, this region of TBP suffices for normal gene regulation and normal growth in at least some strains of yeast (reviewed in ref. 2). Perhaps the TBP of *Arabidopsis* crystallized out so obligingly because it has only 18 amino acids preceding its evolutionarily conserved region.

The structure of TBP has two remarkable features. First, TBP resembles a 'molecular saddle' which straddles the DNA (see figure). The concave underside of the saddle makes extensive contacts with the DNA via a curved eight-stranded antiparallel β -sheet. This DNA-binding surface of TBP is well-defined because mutations in TBP that weaken its binding to DNA or alter its DNA-binding specificity line the concave underside of the molecular saddle¹. This curved antiparallel β -sheet represents a new kind of three-dimensional protein fold for binding DNA, perhaps because, unlike most DNA-binding proteins that bind to DNA in its major groove, TBP interacts with DNA in its minor groove^{3,4} and must function with other DNA sequences for promoters that lack TATA boxes^{5,6}.

Second, nearly all of the conserved region of TBP consists of two topologically identical 88–89 amino-acid domains, whose α -carbon backbones follow the same path in space even though their amino-acid sequences are only 31 per cent identical. As the patterns of functional groups on the surfaces of the two domains are completely different, TBP should still be able to interact asymmetrically with a variety of regulatory proteins. Non-identical amino acids in the two domains should also allow TBP to recognize asymmetric TATA-box sequences (like TATAAA) and provide directionality to transcriptional initiation by RNA polymerase II. But before the precise orientation in which TBP sits on the DNA can be worked out, the solu-

tion of a co-crystal structure with TATA-box-containing DNA will be needed.

The central role of TBP in transcription was emphasized last year with the discovery that TBP is required for transcription by all three nuclear RNA polymerases (reviewed in ref. 7). Con-

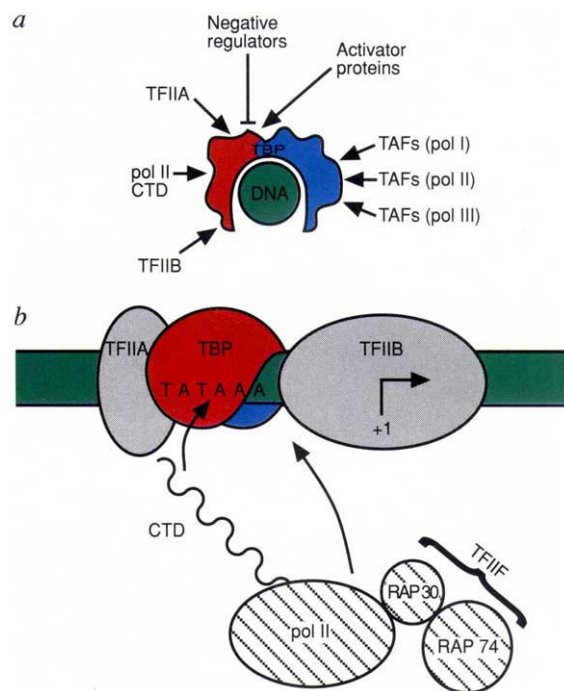
dition are somewhat obscure^{16,17}. Mutational analysis of TBP, guided by knowledge of its three-dimensional structure, should soon reveal precisely how all these factors are aligned on the surface of TBP, which will be invaluable for understanding how transcription is activated and how a promoter selects the appropriate kind of RNA polymerase.

For each of the nuclear RNA polymerases in higher eukaryotes, TBP associates with a unique set of TAFs to form a multisubunit protein complex^{7,9}.

These TAFs probably bind to the conserved region of TBP because the non-conserved amino-terminal domain of TBP is not essential in *Saccharomyces cerevisiae* (reviewed in ref. 2). For RNA polymerases I and III, it seems likely that other promoter-specific DNA-binding proteins recruit the right kind of RNA polymerase by making contact with the appropriate kind of TBP–TAF complex (part *a* of figure). In each case, one of the TAFs should either contact its cognate RNA polymerase or control the ability of another protein to do so. The problem here is that, for RNA polymerases I and III, some of the initiation factors are still not well-defined.

For RNA polymerase II, the situation is quite different. Even though the production of precursors for messenger RNA is a highly regulated process, RNA polymerase II can carry out basal (non-activated) transcription *in vitro* in the absence of TAFs. In this case, binding of TBP to the TATA box allows subsequent binding of the general initiation factors TFIIA and TFIIB to the TBP–promoter complex⁸ (part *b* of figure). TFIIA is probably not necessary for basal transcription *in vitro* in reactions that lack TAFs, but TFIIA binds

directly to TBP^{18,19}, and mutagenesis experiments have shown that TFIIA recognizes a region containing a positively charged α -helix on the convex upper surface of TBP's first conserved domain^{1,20}. Where TFIIB binds on TBP (assuming it does!) is still a mystery, because there are no known TBP mutations that prevent basal transcription *in vitro* without also preventing binding of



a, Transcription factors that interact with the convex outer surface of TBP's 'molecular saddle'. The three nuclear RNA polymerases (I, II and III) apparently function with unique sets of TAFs which bind to TBP^{7,9}. TFIIA and TFIIB are basal factors required for transcription by RNA polymerase II. Competition between activator proteins and negative regulators for binding to TBP may regulate transcription by controlling the association with TBP of TFIIA and TFIIB. Only the binding site for TFIIA on TBP's surface has been well-defined. The two topologically identical domains of TBP are shown in red and blue. *b*, Assembly of a basal transcription complex containing RNA polymerase II. TFIIA, TFIIB and TFIIF are general initiation factors (reviewed in ref. 23). TFIIA, TFIIB and the carboxy-terminal heptapeptide repeat domain (CTD) of the largest subunit of RNA polymerase II all probably bind to TBP^{8,21}. Transcriptional initiation also requires other general initiation factors to associate with a pre-initiation complex containing TFIIA, TFIIB, TFIIF and RNA polymerase II (ref. 23).

sequently, the convex outer surface of TBP's molecular saddle must interact with a large number of other proteins (see part *a* of figure). These include general initiation factors for RNA polymerase II (ref. 8), TBP-associated factors (TAFs; reviewed in refs 7 and 9), repressor and activator proteins that bind TBP^{10–15}, and several yeast proteins (yeast TAFs?) whose roles in transcrip-

TBP to DNA. TFIIB may even recognize an altered DNA structure induced in the promoter by the binding of TBP. A promoter complex containing TBP and TFIIB can recruit RNA polymerase II in conjunction with another general initiation factor, TFIIF (reviewed in ref. 2; part *b* of figure). This association of RNA polymerase II with the pre-initiation complex seems to involve direct binding of the carboxy-terminal heptapeptide-repeat domain (CTD) of its largest subunit to TBP (ref. 21). Where the CTD binds on TBP is also still unknown.

If RNA polymerase II can initiate transcription *in vitro* without DNA-binding transactivator proteins, how does regulation work? As it happens, several kinds of regulators interact directly with TBP (part *a* of figure). First, there exist at least several negative regulators (NC1, NC2, D,1 . . .) that interact with TBP-promoter complexes and prevent binding of TFIIA and TFIIB to the promoter^{10,11}. The existence of these negative regulators is sufficient to explain why at least some activator proteins activate transcription *in vitro* by causing TFIIB to bind to the promoter²². Several activator proteins (such as herpes simplex virus VP16, adenovirus E1A and Epstein-Barr virus Zta) also bind directly to TBP¹²⁻¹⁵. These activator proteins might reverse the effects of the negative regulators on the association of TFIIA and TFIIB with TBP by competing with the negative regulators for binding to the surface of TBP.

Unfortunately, this simple competition

model for transcriptional activation would not easily explain why TAFs are also necessary for transcriptional activation *in vitro*⁹. That is why mutations on the surface of TBP that precisely define the binding sites for TAFs, negative regulators, transactivator proteins and TFIIB are so important. The study of

these mutant forms of TBP should go a long way towards solving the regulation problem. □

Jack Greenblatt is in the Banting and Best Department of Medical Research, University of Toronto, 112 College Street, Toronto, Ontario M5G 1L6, Canada.

OCEAN CIRCULATION

Global warming, ocean cooling

Allyn Clarke

SEA water at intermediate depth in the northeastern North Atlantic is cooler and fresher (less saline) now than in the 1960s, Read and Gould report on page 55 of this issue¹. In a world accustomed to reading about global warming, this result underlines the fact that such environmental changes do not occur smoothly and monotonically over time or in the same manner in all parts of the climate system. And along with related observations, it provides us with a new perspective on the part played by the North Atlantic in climate variability.

This is not the first time cooling of the North Atlantic deep and intermediate waters has been observed. The US Coast Guard occupied hydrographic stations from Ocean Weather Ship Bravo during 1964-74. These showed² a marked freshening and cooling of the upper 400 metres of the Labrador Sea beginning in 1968. At the same time, the waters deeper than 400 metres warmed and became more saline.

The severe winter of 1971-72 ended this pattern. The surface waters were sufficiently cooled to permit deep convection which resulted in a cooling and freshening of the deeper waters and a return to more saline surface waters. Unfortunately, the measurements ceased at the end of 1973 when the weather ships were withdrawn. In the summer of 1981, a separate programme (Transient Tracers in the Ocean) observed that the North Atlantic deep water flowing south along the east Greenland continental slope was much colder and fresher than had previously been observed³.

Each of these earlier observations dealt with variations of the properties of North Atlantic deep water close to their source regions within the Nordic and Labrador seas. The 1991 observations of Read and Gould are unique in what they may reveal about how these colder and fresher waters spread from their source regions into the rest of the ocean interior.

The North Atlantic, under present-day climate, is one of the deep ocean's two principal contacts with the atmosphere, the other being the Southern Ocean. As water flows north into the high-latitude

North Atlantic, it loses heat to the atmosphere and becomes colder and denser. In the Labrador and Nordic seas near the end of winter, the surface waters become cold, saline and dense enough to convect to depths greater than 1,000 metres. These then spread into the rest of the global ocean as its intermediate and deep waters. This process results in a worldwide overturning circulation of which the northward flow of warm saline water in the North Atlantic and the resulting air-sea heat exchange is an important part.

Calculations using simple global ocean models⁴ have shown that the present global circulation pattern with convection in both the North Atlantic and around Antarctica is sensitive to changes in the freshwater fluxes. Increases in the excess of precipitation over evaporation over the North Atlantic can shut down its convection leaving deep convection only around Antarctica. Several ocean modellers are looking at this circulation as the driving mechanism associated with the advance and decline of ice ages.

Dickson *et al.* argued⁵ that the recent cooling and freshening began as an increase in the outflow of low salinity Arctic water in the second half of the 1960s. This advected clockwise around the North Atlantic in the upper waters, suppressing convection through the end of the 1960s. In the winter of 1972, this fresh water began to enter the intermediate waters and was advected back into the eastern basins and the Nordic seas.

Pollard and Pu had previously argued⁶ that the observed decrease in salinity in the upper waters of the eastern North Atlantic during 1967-76 could be accounted for by an increase in excess precipitation over evaporation by 0.16 m yr⁻¹, or about 20-30 per cent. Even with weather ships in the region, we are not in a position to say whether such an increase occurred or not. We still do not have the techniques or instruments to estimate precipitation over the sea. Over the near future, it would appear that the only possible strategy for estimating the freshwater flux in the North Atlantic is through the coupling of an upper-ocean

1. Nikolov, D. B. *et al.* *Nature* **360**, 40-45 (1992).
2. Greenblatt, J. *Cell* **66**, 1067-1070 (1991).
3. Starr, D. B. & Hawley, D. K. *Cell* **67**, 1231-1240 (1991).
4. Lee, D. K., Horikoshi, M. & Roeder, R. G. *Cell* **67**, 1241-1250 (1991).
5. Pugh, B. F. & Tjian, R. *Genes Dev.* **5**, 1935-1945 (1991).
6. Cormack, B. P. & Struhl, K. *Cell* **69**, 685-696 (1992).
7. Sharp, P. A. *Cell* **68**, 819-821 (1992).
8. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. *Cell* **56**, 549-561 (1989).
9. Pugh, B. F. & Tjian, R. *J. Biol. Chem.* **267**, 679-682 (1992).
10. Meisterernst, M. & Roeder, R. G. *Cell* **67**, 557-567 (1991).
11. Inostroza, J. A., Mermelstein, F. H., Ha, I., Lane, W. S. & Reinberg, D. *Cell* **70**, 477-489 (1992).
12. Stringer, K. F., Ingles, C. J. & Greenblatt, J. *Nature* **345**, 783-786 (1990).
13. Horikoshi, N. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 5124-5128 (1991).
14. Lee, W. S., Kao, C. C., Bryant, G. O., Liu, X. & Berk, A. J. *Cell* **67**, 365-376 (1991).
15. Lieberman, P. M. & Berk, A. J. *Genes Dev.* **5**, 2441-2454 (1991).
16. Eisenmann, D. M. *et al.* *Genes Dev.* **6**, 1319-1331 (1992).
17. Koleske, A. J., Buratowski, S., Nonet, M. & Young, R. A. *Cell* **69**, 883-894 (1992).
18. Cortes, P., Flores, O. & Reinberg, D. *Molec. cell. Biol.* **12**, 413-421 (1992).
19. Coulombe, B. *et al.* *Gene Expr.* **2**, 99-110 (1992).
20. Ranish, J. A., Lane, W. S. & Hahn, S. *Science* **255**, 1127-1129 (1992).
21. Ushuva, A. *et al.* *Cell* **69**, 871-881 (1992).
22. Lin, Y. S. & Green, M. R. *Cell* **64**, 971-981 (1991).
23. Zewel, L. & Remberg, D. *Curr. Opin. Cell Biol.* **4**, 488-495 (1992).

model with regular measurements of surface and upper-ocean salinity from merchant vessels and moored or drifting profiling instruments, for which suitable devices must be developed quickly.

Clearly, long-term, high-quality hydrographic measurements are also necessary. The data on which Dickson *et al.* based their analysis were collected by various national fisheries or other research programmes, many of which are no longer running. But annual oceanographic monitoring following each winter's convection is an important component of World Ocean Circulation Experiment (WOCE). CONVEX-91 (a British contribution to WOCE), from which Read *et al.* draw their data, shows that colder, fresher waters are still migrating

through the deeper North Atlantic. The further intensive observations of the North Atlantic planned for 1994–97 under WOCE should provide a new basis for models connecting air–sea fluxes, oceanic convection, North Atlantic variability and the large-scale meridional overturning of the global ocean. □

Allyn Clarke is at the Bedford Institute of Oceanography, PO Box 1006, Dartmouth, Nova Scotia, Canada B2Y 4A2.

1. Read, J. F. & Gould, W. J. *Nature* **360**, 55–57 (1992).
2. Lazier, J. R. N. *Atmos. Ocean* **18**, 227–238 (1980).
3. Brewer, P. G. *et al. Science* **222**, 1237–1239 (1983).
4. Stocker, T. F. & Wright, D. G. *Nature* **351**, 729–732 (1991).
5. Dickson, R. R., Meinke, J., Malmberg S.-A. & Lee, A. J. *Prog. Oceanogr.* **20**, 103–151 (1988).
6. Pollard, R. T. & Pu, S. *Prog. Oceanogr.* **14**, 443 (1985).

CYSTIC FIBROSIS

ATP and chloride conductance

Jeffrey J. Wine and Samuel C. Silverstein

CYSTIC fibrosis results from mutations in the gene for the cystic fibrosis transmembrane conductance regulator (CFTR), a chloride channel that is controlled in complicated and ill-understood ways¹. CFTR is a member of a large family of transmembrane transport proteins that are predicted to have 12 membrane-spanning regions and two nucleotide-binding domains; it also uniquely possesses a large, highly charged cytosolic region containing numerous phosphorylation consensus sites. In accord with predictions from structural data, opening of CFTR channels requires both phosphorylation and binding of ATP².

Why is the regulation of this channel so complex? On page 79 of this issue³, Quinton and Reddy report that in human sweat-duct cells activated by cyclic AMP, nonhydrolysable ATP analogues can stimulate chloride conductance via CFTR and that intracellular ATP levels of 5 mM are needed to maintain normal levels of chloride conductance. They put forward a new and important idea — that CFTR is sensing either the energy charge of the cell [(ATP + 1/2 ADP)/(ATP + ADP + AMP)] or the cellular ATP concentration itself, and that one or both of these factors regulates ion transport by CFTR.

Human sweat ducts are tiny organs which are difficult to study. Their usefulness is that, of any tissue, they express among the highest levels of CFTR on their luminal surface and have one of the highest chloride conductances; such conductance is absent in ducts afflicted with cystic fibrosis⁴. By permeabilizing the basolateral membrane of the duct with a bacterial toxin that creates large pores, the cytosolic levels of nucleotides can be

altered and the chloride conductance of the luminal membrane measured. Under these conditions, Quinton and Reddy found that levels of ATP many times higher than the K_m for activating the cAMP-dependent protein kinase PKA did not produce appreciable increases in conductance. But after about 1 mM of ATP, which alone did not increase chloride conductance, the conductance rose steeply up to 10 mM of ATP, the highest concentration used. With the ATP level held at 0.5–1.0 mM to maintain phosphorylation of CFTR, large increases in conductance were produced by 5 mM additions of any four different non- or poorly hydrolysable analogues of ATP, whereas AMP was inhibitory.

These results are surprising. Previous experiments with excised patches expressing recombinant CFTR suggested that opening of CFTR in such patches requires hydrolysis of ATP, and concentrations for half-maximal activation were only about 260 μ M ATP². There are, however, several differences in the conditions under which the two experiments were carried out. CFTR was endogenous to the ducts but recombinant in the excised-patch experiments, and magnesium, an essential cofactor for ATP hydrolysis, was held at different concentrations. Moreover, bath chloride was replaced with gluconate in the duct experiments, and whereas pores formed by the bacterial toxin should allow free diffusion of the ATP analogue used, the duct apparently retained most of its metabolic machinery and cytosolic structure so that actual concentrations of nucleotides at the cytosolic face of the apical membrane may differ from those in the bath. For example, bath ADP, which inhibited

CFTR opening in excised patches⁵, stimulated opening in the duct. This was presumably because the permeabilized duct retains adenylate kinase, and thus the ability to form ATP and ADP. Moreover, the intact duct cells retain phosphatases capable of dephosphorylating the regulatory domain of CFTR. Such phosphatases are presumably less readily available in excised patches. Clearly, all these differences must be addressed experimentally before the results can be directly compared.

Taking Quinton and Reddy's observations at face value, it seems difficult to reconcile them with a mechanism that depends purely upon hydrolysis. Here, other recent experiments that hint at complexities in the ATP-regulation of CFTR and related channels come to mind. If ADP (and AMP?) directly inhibit CFTR channel opening⁵, and if ADP and AMP are present at significant concentrations in the permeabilized duct, increased levels of ATP will be needed for conductance. Some clues to what is going on might also be provided from studies of P-glycoprotein (MDR1), an integral membrane protein related to CFTR, which confers multi-drug resistance upon cells in which it is expressed⁶. P-glycoprotein can apparently function either as a chloride channel activated by swelling, or as a transporter of a wide range of lipid-soluble compounds⁷. Opening of the P-glycoprotein channel is stimulated by allosteric ATP binding⁸, but transport requires ATP hydrolysis⁹. The protein cannot simultaneously carry out both functions⁸.

Quinton and Reddy's experiments will mean that even more attention is focused on mechanisms of CFTR regulation, on whether CFTR has several functions, and, if it does, whether they are distinct and use ATP differently. Numerous efforts are now underway to provide rational treatments for cystic fibrosis, and these issues are of great relevance in unravelling the pathophysiology of the disease. □

Jeffrey J. Wine is in the Cystic Fibrosis Research Laboratory, Stanford University, Stanford, California 94305, USA; Samuel C. Silverstein is in the Department of Physiology and Cellular Biophysics, Columbia University College of Physicians and Surgeons, Rm 11-511, 630 West 168th Street, New York, New York 10032, USA.

1. Riordan, J. R. *et al. Science* **245**, 1066–1073 (1989).
2. Anderson, M. P. *et al. Cell* **67**, 775–784 (1991).
3. Quinton, P. M. & Reddy, M. M. *Nature* **360**, 79 (1992).
4. Quinton, P. M. *Nature* **301**, 421–422 (1983).
5. Anderson, M. P. & Welsh, M. J. *Science* **257**, 1701–1704 (1992).
6. Bradley, G., Juranka, P. F. & Ling, V. *Biochim. Biophys. Acta* **948**, 87–128 (1988).
7. Valverde, M. A. *et al. Nature* **355**, 830–833 (1992).
8. Gill, D. R. *et al. Cell* **71**, 23–32 (1992).
9. Ambudkar, S. V. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 8472–8476 (1992).

Rubbish birds are poisonous

Jared M. Diamond

MANY animal species, of which monarch butterflies and poison-dart frogs are the best known, contain poisons and bitter-tasting chemicals that serve to deter predators. But not birds — or so it had seemed, for in the latest issue of *Science*¹ Dumbacher and colleagues describe three species that are not only poisonous but highly so. The discovery provides an astonishing example of convergent evolution, and also deserves study for possible geographical variation in the birds' poison, for their involvement in an extended mimicry complex and for how the phenomenon bears on the evolutionary history of birds of paradise.

The birds concerned are species of the genus *Pitohui*, common medium-sized omnivores that are old endemics of the island of New Guinea. When Dumbacher and colleagues collected and handled a pitohui, they noticed a strong sour smell and felt numbness and a burning sensation of the mouth. Natives of New Guinea branded the Hooded Pitohui as a 'rubbish bird' because it has a bitter taste and is unfit for eating unless skinned and carefully prepared. Ethanol extracts yielded the steroidal alkaloid homobatrachotoxin, a nerve and muscle poison previously known only as an active principle of poison-dart frogs (skins of the frogs are used by Amerindians to poison their blowgun darts). Injection into mice of extracts of pitohuis (especially of the skin and feathers, but also of skeletal muscle) caused paralysis of the hind limbs and prostration, leading at 'high' doses (estimated at 0.05 µg of homobatrachotoxin) to convulsions and death in as little as 15 minutes. How pitohuis make or acquire the poison, and how their own nerves and muscles survive it, remain unknown.

The presumed function of the poison is to repel predators, an interpretation which is reinforced by the pitohuis' strong smell which would serve as a warning signal. Just as a skunk's striking black-and-white colour constitutes a warning signal distinct from its stink, the pitohuis' stink must be due to a separate chemical; homobatrachotoxin itself is odourless.

Pitohuis were already notorious in ornithology because one of them, the Variable Pitohui, exhibits the most striking geographical variation in plumage of any bird species found in New Guinea^{2,3}. Some Variable Pitohui races have a black-and-maroon plumage like the Hooded Pitohui; others have a uniformly dull plumage like the Rusty Pitohui; and still others have different plumage patterns. All three of these species proved

poisonous, the Hooded Pitohui being the most poisonous. The observations raise the question of possible Müllerian mimicry (the phenomenon of several species sharing similar warning signals or poisons), and of whether poison concentrations in a given pitohui species vary geographically. Dumbacher *et al.* obtained their Hooded Pitohuis from New Guinea's Sogeri Plateau, and natives of a second locality (the Kaironk Valley) have also been reported as

predator to pick out any particular individual. I had assumed that the far-reaching convergence among all these species served either for predator avoidance or for flock cohesion promoted by shared social signals. But with the discovery that at least three of the leader species are poisonous, the question arises whether the leading babbler is also poisonous (Müllerian mimicry), and whether the follower species' similar plumages stem in part from Batesian mimicry.

The existence of poisons in pitohuis may be relevant to New Guinea's most famous old endemics, the birds of paradise, whose males evolved through sex-

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Bitter bird — a Hooded Pitohui, *Pitohui dichrous*. (Courtesy of William S. Peckover.)

saying that the species tastes bitter⁴. In other areas of New Guinea, though, my field associates and I dissected dozens of pitohuis, and my associates ate them, without untoward effects. Either I missed a great discovery, or else only some populations are poisonous.

Mimicry may extend phylogenetically far beyond pitohuis and may include Batesian mimicry (nonpoisonous species mimicking poisonous ones), because pitohuis are also notable as nuclear species of New Guinea's "brown and black flocks"⁵. These itinerant foraging flocks consist of about 35 medium-sized bird species, belonging to seven families, which wander in search of food. The flocks are led by a babbler and by one or more of five pitohui species, including the three now known to be poisonous. The plumage of all members of these flocks is brown or black or both. As the flocks go past, the whirl of brown and black makes it hard for a would-be

ual selection to be the most bizarrely plumed birds in the world. Despite great species differences in plumage among the males, females are much more conservative in their brown and black plumage. At least 15 species of bird of paradise join the pitohui flocks, and the taste of birds of paradise has been reported as "the most shocking flesh I have ever eaten . . . as bitter as gall . . . truly abominable"⁶. So interactions with poisonous brown and black birds may have been a long-standing selective force on birds of paradise.

Finally, there is a broader moral to this tale. Pitohuis have been known to

1. Dumbacher, J. P. *et al.* *Science* **258**, 799–801 (1992).
2. Beehler, B. M. *et al.* *Birds of New Guinea* Plate 42 (Princeton University Press, 1986).
3. Rand, A. L. & Gilliard, E. T. *Handbook of New Guinea Birds* 435–436 (Weidenfeld & Nicolson, London, 1967).
4. Majnep, I. S. & Bulmer, R. *Birds of My Kalam Country* 103 (Oxford University Press, 1977).
5. Diamond, J. M. *Emu* **87**, 201–211 (1987).
6. Pratt, A. E. *Two Years Among New Guinea Cannibals* 219, 220, 258 (Seeley, London, 1906).

Western scientists since 1827, and to New Guineans for at least 40,000 years. Thousands of specimens are scattered among the world's leading zoological museums. The birds are common and are observed every year by dozens of bird-watchers visiting New Guinea. Yet this remarkable example of chemical defence, convergent evolution and possible mimicry escaped scientific attention until now. Like most other New Guinea bird species, pitohuis are confined to rainforests, which (as elsewhere through-

out the humid tropics) are shrinking under the onslaught of commercial export logging. What other unsuspected wonders are disappearing in each hectare of rainforest felled? What other treasures of biological knowledge are becoming lost with the rapid acculturation of the world's few remaining Stone Age hunters? □

Jared M. Diamond is in the Department of Physiology, University of California Medical School, Los Angeles, California 90024, USA.

INTERSTELLAR DIAMONDS

Rich pickings for astronomers

I. P. Wright

DIAMONDS have been found in dust clouds surrounding forming stars, by L. J. Allamandola *et al.*¹. The observations will be of great interest to meteoriticists, who have also found diamonds in primitive, unshocked meteorite samples. According to the laboratory studies, diamonds constitute a small fraction, roughly one to five per cent, of the total carbon meteorites². It is probable that this represents a lower limit to the abundance of diamonds present in the protostellar cloud that collapsed to form our Solar System².

The possibility that diamonds could exist in the interstellar medium was first expounded by Saslaw and Gaustad³, who argued that material of this nature might explain certain features in the infrared spectrum of dust. The same authors were also bold enough to suggest that studies of meteorites may be relevant to this issue, a suggestion undoubtedly inspired by a previous debate on the origin of diamonds in meteorites like the Canyon Diablo iron, which was responsible for excavating Meteor Crater in Arizona. Although some proponents considered that diamonds in meteorites had resulted from high static pressures, as would be obtained inside a meteorite source body of planetary dimensions, it was subsequently shown that they were more likely to have been formed by shock-transformation of graphite during the impact event⁴.

Diamonds with characteristics that could be ascribed to surviving interstellar dust grains were eventually isolated from primitive chondritic meteorites by Lewis *et al.*² in 1987. These nanometre-sized crystallites have since become the subject of intensive laboratory investigation; of fundamental interest is their origin. In this regard, it may have been anticipated that measurement of the stable carbon-isotope composition of the diamonds would have been informative. Ironically, however, this parameter is not substan-

tially different to average Solar System material, although it does show some small but significant variability⁵.

Of more use are isotope measurements of the xenon trapped in the diamonds which show an enrichment in both heavy and light isotopes, commensurate with a supernova origin for this component^{6,7}. An enrichment of other products of this rapid-process (r-process) nucleosynthesis has also been detected in the case of barium, although apparently not for strontium⁸.

Detection of these nuclides is formidable, as they are present at the level of one atom per 10^4 – 10^5 crystallites. Current wisdom holds that the diamond crystallites may have been formed by condensation, perhaps via a process analogous to chemical vapour deposition, in the outflow from a carbon star; the isotopically anomalous trace constituents were then implanted as ions expelled from a subsequent supernova explosion⁷. Equally well, however, it has been argued that the diamonds are in fact the result of shock-transformation of interstellar graphite grains⁹.

But were they formed in a single event or in many? The size distribution (log-normal) is consistent with growth rather than fragmentation (which gives a power law), taken to imply that the diamond population represents grains in an early stage of evolution⁷ and so, by extension, from a single event. Nitrogen, present at levels of 0.2–1.3 per cent⁵ (the high concentration could indicate the presence of unidentified nitrides), reveals an isotope composition that is remarkably uniform and quite different from the average for the Solar System. The constancy of the isotopic composition, together with the range in nitrogen contents, tends to indicate a single source for the diamonds (or at most two, one with nitrogen and the other without).

Nonetheless, there is the isotopic car-

bon variation to be accounted for. And there are now known to be some diamonds present in primitive meteorites that were apparently formed within the solar nebula itself¹⁰. So although the case for diamond growth by condensation in stellar outflows or through shock processing of interstellar graphite is strong, it is apparent that diamonds can also form in the disks around forming stars.

Allamandola *et al.*, who have just reported their results in the latest *Astrophysical Journal*¹, obtained spectra of four dense molecular clouds, each known to contain high levels of dust, around putative protostars. In each case, they see a previously unobserved absorption feature at a wavelength of 3.47 micrometres ($2,880\text{ cm}^{-1}$). Although this is in the spectral region where C–H stretching vibrations occur, the absorption is not due to $-\text{CH}_3$ or $-\text{CH}_2-$ groups, so that organic molecules are not involved. Instead the authors suggest that a diamond-like structure is responsible, the new feature being due to hydrogen present on its surface. If so, a few per cent of the cosmic complement of carbon is in this form — close to the value suggested from meteorite studies.

Sadly, even if diamond crystallites are common in protostellar environments, this does not answer the question about their origin. To this end it is interesting to note that Colangeli *et al.*¹¹ have shown that diamonds synthesized on Earth by detonation mechanisms can yield subtle differences in their infrared absorption characteristics. Furthermore, only in certain cases do the spectral data from these samples match what is observed in the diamonds from meteorites. Clearly ever higher-resolution spectral analyses of circumstellar dust, meteoritic diamonds and terrestrial synthetic analogues will be one way towards unravelling the origin and evolution of interstellar diamonds. □

I. P. Wright is in the Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

- Allamandola, L. J., Sandford, S. A., Tielens, A. G. G. M. & Herbst, T. M. *Astrophys. J.* **399**, 134–146 (1992).
- Lewis, R. S., Tang, M., Wacker, J. F., Anders, E. & Steel, E. *Nature* **326**, 160–162 (1987).
- Saslaw, W. C. & Gaustad, J. E. *Nature* **221**, 160–162 (1969).
- Lipschutz, M. E. & Anders, E. *Geochim. cosmochim. Acta* **24**, 83–105 (1961).
- Russell, S. S., Arden, J. W. & Pillinger, C. T. *Science* **254**, 1188–1191 (1991).
- Manuel, O. K., Hennecke, E. W. & Sabu, D. D. *Nature* **240**, 99–101 (1972).
- Lewis, R. S., Anders, E. & Draine, B. T. *Nature* **339**, 117–121 (1989).
- Lewis, R. S., Huss, G. R. & Lugmair, G. *Lunar Planet. Sci.* **XXII**, 807–808 (1991).
- Tielens, A. G. G. M., Seab, C. G., Hollenbach, D. J. & McKee, C. F. *Astrophys. J.* **319**, L109–L113 (1987).
- Russell, S. S., Pillinger, C. T., Arden, J. W., Lee, M. R. & Ott, U. *Science* **256**, 206–209 (1992).
- Colangeli, L. *et al. Meteoritics* **27**, 210–211 (1992).

Struggling with hieroglyphics

Leo Blitz

It appears that astronomers are witnessing for the first time the birth of an entire galaxy similar to the Milky Way. The gestation is largely hidden from the view of optical telescopes, and is straining the limits of what is measurable with current instruments. In a series of six articles published in the 20 September and 10 October issues of the *Astrophysical Journal*¹⁻⁶, separate groups of European, American and Japanese astronomers have been struggling to interpret observations of an extraordinary object that defies conventional understanding.

Last year, Rowan-Robinson *et al.*⁷ announced the discovery of a galaxy brighter than any known object in the Universe. Although the galaxy, known by its catalogue number F10214+4724, emits about 1,000 times as much radiation as the entire Milky Way, it is optically inconspicuous; it was initially identified at infrared wavelengths. Shortly after the discovery, Brown and vanden Bout⁸ looked for emission from the CO molecule and found it. CO is a tracer of the much more abundant H₂ molecule, and although the initial estimates of the mass of H₂ were too high⁹, the mass of molecular gas alone in F10214+4724 was found to be about 10¹² solar masses, more than the total mass of the Milky Way (which contains only about 10⁹ solar masses of H₂).

Because molecular hydrogen is the raw material from which new stars form, and because the mass of gas alone is more than enough to form an entire galaxy, these two sets of observations strongly suggested that F10214+4724 is a *primaeval* galaxy, one still in the initial stages of formation. The luminosity of the galaxy is probably due to an early generation of massive stars still enveloped in their nascent dust cocoon. The intense visible and ultraviolet stellar radiation would be largely absorbed by the surrounding dust (itself from yet an earlier generation of star formation) and reradiated in the infrared.

The initial excitement was followed by a dose of healthy scepticism. The high luminosity of the galaxy is inferred from its infrared brightness and from the large redshift (implying great distance) of 2.236 measured for its proposed optical counterpart. How do we know that the infrared source and the optical galaxy are the same? The positional uncertainty of the infrared source is one minute of arc, and there are seven optical sources that could plausibly be responsible. Most, but not all, have been ruled out, but there is not yet a published con-

firmation that the galaxy which has the large optical redshift is the infrared source.

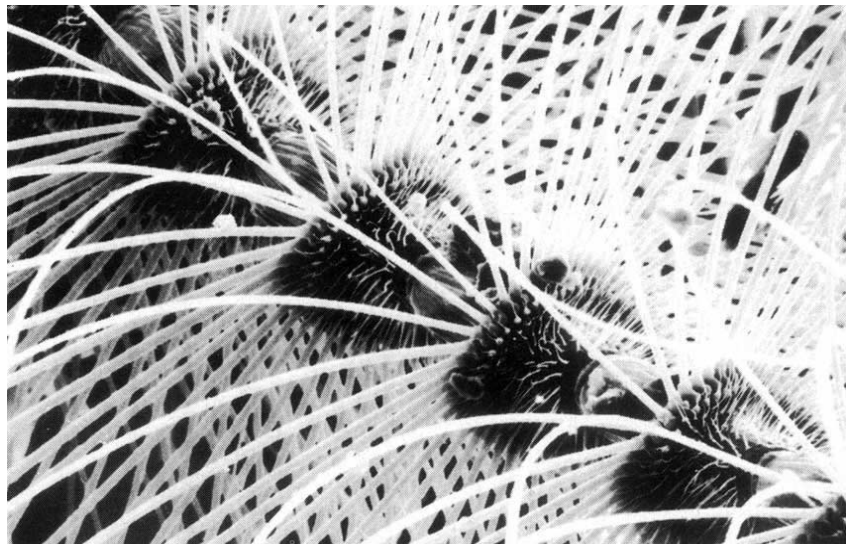
Serious questions have also been raised about the detection of CO. Although the detection itself looks fairly convincing, the width of the observed spectral line is almost unbelievably narrow. The width is determined by the total mass of the galaxy and the inclination of the gas orbits to our line of sight. How could a galaxy with that much mass (the initial mass estimates were 10–100 times that of the Milky Way) have a line much narrower than would the Milky Way or other galaxies of similar mass? It is not clear that a *primaeval* galaxy would have developed into a disk, no less be highly inclined to the line of sight.

This is where the new observations play a conclusive role in verifying the identity and uniqueness of F10214+4724, as well as imposing some new difficulties. Aperture-synthesis observations with the Nobeyama array^{3,4}, a technique fundamentally different from Brown and vanden Bout's single-dish observations, greatly improved the resolution of the CO sighting and show that the emission is real and is centred within 1 arcsecond

of the galaxy identified by Rowan-Robinson *et al.* Yet, the spectral line detected is even narrower than that of the original CO detection. The galaxy would still have to be nearly face on, deviating at most by 8–18° from the line of sight. Furthermore, the Japanese did not detect the highest-velocity components of Brown and vanden Bout's spectrum suggesting that either this gas is very extended, covering more than about 30 arcseconds, or the original spectrum is corrupted.

Brown and vanden Bout's more recent work² argues that the CO emission is indeed extended, because the line is barely brighter when observed with a larger telescope, the 30-m IRAM telescope in Spain, than with the 12-m telescope in Arizona that they used initially. However, if the galaxy is as large as they require, then it would have an extent of at least 100–200 kiloparsecs; the galaxy would be uniquely large. For comparison, the Milky Way (itself rather macho as spiral galaxies go), has a diameter of about 40 kiloparsecs. But the flux density (a measure of the surface brightness) recorded at the Nobeyama array implies that the Japanese are observing all of the emission seen with the IRAM telescope, and that all of the emission detected at IRAM is concentrated within a radius comparable to the size of the Milky Way. That is, the galaxy is not uniquely large. Thus it

Tuned into good vibrations



MALE mosquitoes, unlike females, boast large and elaborate antennae which aid them in their quest for mates. The antennae also helped Gregory Paulson of Washington State University to gain the overall first prize in this year's International Photomicrography Competition sponsored by Polaroid. The familiar whining sound of female mosquitoes in flight reflects the wing-beat frequency of the species and causes the antennae of males of the same species to resonate. Directional cues are detected by massive sense organs at the base of the antennae. Paulson's micrograph, reproduced here, shows the rings of fibrils encircling the shaft of a male's antenna (magnification, $\times 340$).

P.T.

appears either that there is an instrumental problem and the highest velocity gas is not real (or possibly, that calibration uncertainties have been underestimated), or that some of the molecular gas is concentrated in a galaxy-sized part, and another part, much more extended and of nearly equal mass, does not emit visible radiation.

Observations at IRAM by a largely European group suggest that the difference between the properties of F10214+4724 and other known galaxies may not be as extreme as was first thought. The dust mass⁵ is somewhat less than originally determined by Rowan-Robinson *et al.*; and neither is there as much H₂ as others estimated⁶. The IRAM group concludes that F10214+4724 is similar to but at the extreme end of a class of ultraluminous galaxies identified from the IRAS infrared galaxy survey. These galaxies are as luminous as quasars, but like F10214+4724, emit most of their radiation in the infrared portion of the spectrum rather than in the optical range. Their luminosities are generally attributed to an enormous burst of star formation triggered by the merger of two spiral galaxies. The authors argue that F10214+4724 appears to be a more extreme example of these exotic, but better known galaxies.

But once again we run into observational and conceptual difficulties. The flux density in the CO line measured by the European investigators is considerably less than that measured by the Japanese team, even though their single-dish observations measured a larger area than did the Nobeyama array. The opposite should have been observed. In any event, a fundamental conundrum appears in trying to fathom all the data. The lower mass estimate makes the narrow line widths more palatable, and does not require an unlikely viewing angle. On the other hand, the lower mass estimate is difficult to reconcile with the observed carbon abundance.

Neutral gaseous carbon was detected¹ from F10214+4724 in two different transitions, which makes it possible to get a reasonably good estimate of the total mass of atomic carbon, about $1-3 \times 10^8$ solar masses. Carbon is believed to be a product of stellar nucleosynthesis, its presence in the interstellar medium signalling its manufacture in the centres of massive stars which then return it to the interstellar medium when the stars explode as supernovae. Near the Sun, the carbon abundance relative to hydrogen is 4×10^{-3} by mass, but this has been built up over the age of the Universe. In F10214+4724, if the H₂ mass is as small as 1×10^{11} solar masses as suggested by the European and the Japanese investigators, then the carbon abundance

would be nearly equal to the solar abundance. How is it possible for a galaxy five times younger than the Milky Way to have manufactured and dispersed such a disproportionate amount of carbon? If Brown and vanden Bout's H₂ mass estimate of about 10^{12} masses is closer to the correct one, the carbon abundance is more palatable, but the discrepancy between the large mass and narrow line width is once more resurrected. Does it make sense that we observe a unique galaxy, which could be oriented in any direction, at a highly improbable angle? Or that a massive primaeval galaxy should be a thin disk?

As observers wrestle with these questions, other tantalizing questions and possibilities arise. For example, in normal galaxies, massive stars make up only a small fraction of the total stellar inventory of a galaxy. Does F10214+4724 have its normal complement of low-mass stars? According to standard stellar evolutionary theory, the large burst of massive-star formation needed to explain the infrared luminosity would leave enormous numbers of black hole remnants within a few million years of the starburst that we are currently observing. So, unless the galaxy is unique (an unlikely prospect) there may be a large number of galaxies with tens of millions of black holes in their midst.

In spite of the inconsistencies, it is important to recognize that the problems in deciphering the observations of F10214+4724 do not come from the shortcomings of the observers, but are largely the result of instruments being pushed to their limits. Notwithstanding all the difficulties, one thing is clear: F10214+4724 is an extraordinary galaxy in which a uniquely large fraction of its mass is in the form of molecular gas making massive stars at a prodigious rate. The observational evidence does suggest that it is a galaxy that is in an early evolutionary stage and F10214+4724 may thus be a Rosetta Stone for understanding how galaxies form. However, it will take much more work to decipher its hieroglyphics. □

Leo Blitz is in the Astronomy Department, University of Maryland, College Park, Maryland 20742, USA.

1. Brown, R. L. & vanden Bout, P. A. *Astrophys. J.* **397**, L11-L14 (1992).
2. Brown, R. L. & vanden Bout, P. A. *Astrophys. J.* **397**, L19-L22 (1992).
3. Kawabe, R., Sakamoto, K., Ishizuki, S. & Ishiguro, M. *Astrophys. J.* **397**, L23-L26 (1992).
4. Sakamoto, K., Ishizuki, S., Kawabe, R. & Ishiguro, M. *Astrophys. J.* **397**, L27-L30 (1992).
5. Downes, D. *et al.* *Astrophys. J.* **398**, L25-L28 (1992).
6. Solomon, P. M., Downes, D. & Radford, S. J. E. *Astrophys. J.* **398**, L28-L32 (1992).
7. Rowan-Robinson, M. *et al.* *Nature* **351**, 719-721 (1991).
8. Brown, R. L. & vanden Bout, P. A. *Astr. J.* **102**, 1956-1959 (1991).
9. Solomon, P. M., Radford, S. J. E. & Downes, D. *Nature* **356**, 318-319 (1992).

DAEDALUS

Fulminating soot

In the past few years, the humble carbon arc has proved a prolific source of wonderful new carbon allotropes — buckminsterfullerene, other fullerenes and carbon nanotubes. All these molecules assemble themselves from carbon evaporated from the electrodes. Now Ostwald's rule says that an element condenses from the vapour into its least stable allotropic form. Thus phosphorus vapour condenses to yellow phosphorus, not the more stable red variety. So Daedalus reckons that the vapour from a carbon arc should generate the most unstable carbon allotropes of them all. These are the carbynes, linear high polymers of carbon atoms linked by alternate single and acetylenic bonds, and usually too explosive to isolate. But intrepid DREADCO chemists are striking carbon arcs under liquefied noble gases, which should safely quench and contain the growing carbyne chains as they form.

To isolate the products safely, Daedalus will tip in a large excess of some inert polymer. The carbyne will be diluted to harmlessness in the resulting polymeric matrix. The mixed polymer could even be moulded and stretched to align the straight-chain carbyne molecules from end to end of the sample.

At first, Daedalus felt that the product would be a novel organic conductor. A carbyne molecule, as the ultimate fine carbon filament, should conduct electricity. But he then realized the disruptive consequences. Electricity is a flow of electrons. A flow of electrons along a carbyne molecule would shift all the bonding electrons one step along. The molecule would be disrupted into pairs of C₂ radicals in a single-molecule explosion. The radicals would probably then combine with their polymer matrix.

An explosive fired directly by electricity, and going to a solid product, should be very useful. It would react instantaneously throughout the current-carrying volume: no shock-wave would be needed to carry the explosion from point to point. So a plastic containing very little carbyne, far too dilute to propagate a shock-wave and therefore quite safe against flame or impact, could be exploded electrically. It would give an equally dilute blast — simply a sudden expansion of the plastic block.

DREADCO's 'Carbynite' will transform engineering. It will be formed into all sorts of simple, reliable, pre-calibrated, electrically triggered actuators, releasers and jacks. It will split rock, shift jammed loads, drive nails, and operate emergency brakes and stop-valves. Wherever a controlled force needs to be applied once only, safely and certainly, Carbynite will be there.

David Jones

Breslow replies to criticism

SIR — I have just read the complaint by Menger and Haim (*Nature* 359, 666–668; 1992) that they were poorly treated by *J. Am. chem. Soc. (JACS)* in their attempt to criticize some work we published a few years ago. The situation is not as they represent it. Menger's paper was rejected by *JACS* because of referees' reports that it lacked merit; I was not a referee, but the referees saw my comments. Among other flaws, it misrepresented what we had published, then attacked the misrepresentation. It was eventually published in *J. org. Chem. (JOC)*, leading a *JACS* editor to demand in writing to the *JOC* editor that the Menger paper be retracted because of its flaws, particularly the misrepresentations.

The main complaint in Menger's paper was that we had used the phrase "negative experimental catalytic rate constant" to describe our observation that one of the reactions slows as more catalyst is added. The phenomenon is not unusual (it is related to the well-known common ion effect in solvolysis) and is completely confirmed by our later work. It is seen because an intermediate partitions along two different pathways; thus speeding up one path with the catalyst slows the rate of formation of the other product. This proves a common intermediate for the two processes. Even though one would certainly refer to the coefficient of a catalyst as an "experimental catalytic rate constant" for any process that speeds up when catalyst is added, some chemists object to calling it an experimental rate constant when it is a negative number. This is a matter of taste, not an error. The original data, and our new work, fully confirm the fact of the rate decrease. Of course we never invoked negative rates, simply a negative catalytic coefficient to describe quantitatively the slowing of the reaction.

The Haim paper is apparently to be published in *JACS*. It was originally rejected after referees' comments, but eventually accepted. It claims that our original data do not fit very well the expected kinetics for the mechanism we propose, but does not propose an alternative mechanism.

Since that time we have done new work, currently being refereed by *JACS*. Haim and Menger have both seen this new paper. It completely confirms our previous conclusions, including the negative catalytic effect, and shows that a sophisticated kinetic treatment fully accommodates the previous data as well. It also calls attention to our first published work on this system which had many carefully buffered data points. With this foundation, fewer data were

needed for the later work that has been criticized as being incomplete. The new extensive work and the sophisticated kinetic treatment have been presented before experts at meetings such as the reaction mechanism symposium, and have met no objections. The scientific facts fully support our previous conclusions.

Of course science must be self-correcting, but criticism should be well founded. There is no excuse for distorting what has been published in order to attack it. The happy ending to this story is that we were stimulated to do our new studies by outrage over the misrepresentations. This new work has added considerably to our understanding while confirming the previous conclusions. The overall result has let us reinterpret the mechanism of an important enzyme, a reinterpretation supported by others, and it has guided us to the design of an improved enzyme mimic.

Ronald Breslow

Department of Chemistry,
Columbia University,
New York, New York 10027, USA

Coley's toxins

SIR — Brouckaert *et al.*¹, in response to my Commentary², discuss several issues that need to be addressed. First, the optimistic clinical results mentioned by Brouckaert *et al.*, while very impressive and important in their own right, all involve some form of direct, local administration of cytokines within or near the site of the growing tumour. The ability to make such a delivery was a luxury not shared by most of the inoperable soft-tissue sarcoma patients treated by Coley and his contemporaries, nor would it be shared by those patients with similar disease today. It is also not necessary for the successful treatment of a truly sensitive murine tumour (such as Meth A).

Second, Brouckaert *et al.* assert that "... other animal studies in which TNF was used in conjunction with interferons showed a similar beneficial effect for melanomas and for carcinomas." I question the phrase "similar beneficial effect", when the results of those studies are compared with what can be achieved in the treatment of a truly sensitive tumour growing in a normal, immunocompetent syngeneic recipient. A properly timed, single-shot systemic treatment is generally sufficient to give long-term cures in 70–100% of the animals treated. Among the essential prerequisites for this phenomenon is the fact that the tumour must be immunogenic as classically defined^{3,4}. As I implied², the

embryonic origin of the tumour may have no more importance than simply to help identify those tumours which have the highest probability of being immunogenic.

The clinical problem, of course, is that we have no comparable method of establishing the degree of immunogenicity of human tumours *a priori*, as we do with inbred mice. With humans, we must rather deal with a retrospective assessment, the essence of which has been provided for us in the medical records of Coley and his contemporaries. I believe that these records have already given us a perspective on which human tumours are likely to be immunogenic and which are not. Unfortunately, patient selection based on this perspective is decidedly imperfect; most inoperable mesodermally derived tumours were not known to have been cured by Coley's treatment (by the most optimistic estimate, only one in five). Matters are further complicated by the fact that even incurable, non-immunogenic murine tumours can be made to undergo some degree of necrosis in association with a therapeutic response that is only temporary^{3,4}; Coley had made similar observations in the treatment of some human carcinomas. In contrast, because of the admirable follow-up efforts made by Coley's daughter Helen Coley Nauts, we know that a substantial number of inoperable soft-tissue sarcoma and lymphoma patients treated by her father and his contemporaries survived with no evidence of disease for more than 20 years (in my estimation, more analogous to the Meth A model). These records have given us a unique perspective that is essential in guiding us towards proper patient selection.

Finally, I specifically stated that "... serious consideration should be given to a return to an aggressive use of the vaccine ..."; not TNF-related or toxin-related therapy, but specifically the Coley vaccine itself. This is a point made in deference to the fact that the clinical accomplishments of Coley and his contemporaries were much beyond what we would be able to offer these same patients today. If we are going to make changes or improvisations in treatment, this should not be done until after we have at least managed to reproduce the original, basic observations.

Charlie O. Starnes

Amgen, Inc.,
1840 Dehavilland Drive,
Thousand Oaks,
California 91320-1789, USA

1. Brouckaert, P. G. G., Fiers, W. & Lejeune, F. J. *Nature* 358, 630 (1992).

2. Starnes, C. O. *Nature* 357, 11–12 (1992).

3. Berendt, M. J. *et al. J. exp. Med.* 148, 1550–1559 (1978).

4. Berendt, M. J. *et al. J. exp. Med.* 148, 1560–1569 (1978).

Protein crystals go bang

SIR — Daedalus¹ indicates the destructive potential contained in the crystal lattice of antimony and the plans of DREADCO's chemists to produce explosive, yet stable, forms of crystals modified by proteins. But he may be interested to hear that crystallographers have already observed in pure protein crystals some ability to undergo this violent self-destruction.

Like the explosive form of antimony, a slight mechanical impact with a needle often results in rapid shattering and destruction of these protein samples. The small size of protein crystals normally available for X-ray analysis (about 0.1–1 mm³) may have saved many unwary scientists from serious injury! This type of detonation can also be catalysed by heavy metals. Combining 5 mM Hg(NO₃)₂ with crystals of the avian pancreatic polypeptide, for example, results in an impressive and rapid disintegration^{2,3}.

The development of proteins engineered for a highly explosive transition

of crystal habit could be attempted by recombinant DNA techniques and site-directed mutagenesis. Engineered surface loops trapped in a crystal lattice but highly mobile in the non-crystalline state are likely to act as triggers for rapid ignition. Much like the development of gas centrifugation used previously to purify uranium, massive crystallization centrifuges would have to be constructed to produce the extremely large protein crystals required to reach a critical mass⁴. Even the technology for producing a self-destroying 'plastic explosive' bag suggested by Daedalus is available using protein crosslinking agents like glutaraldehyde. This exciting prospect opens up a whole new avenue for generous research funding on the development of peaceful uses for this terrifyingly destructive material.

Jim E. Pitts

Department of Crystallography,
Birkbeck College,
University of London,
Malet Street, London WC1E 7HX, UK

1. Jones, D. *Nature* **358**, 546 (1992).
2. Pitts, J. E. thesis (Univ. Sussex, 1980).
3. Blundell, T. L., Pitts, J. E., Tickle, I. J., Wood, S. P. & Wu, W.-C. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4175–4179 (1981).
4. Pitts, J. E. *Nature* **355**, 117 (1992).

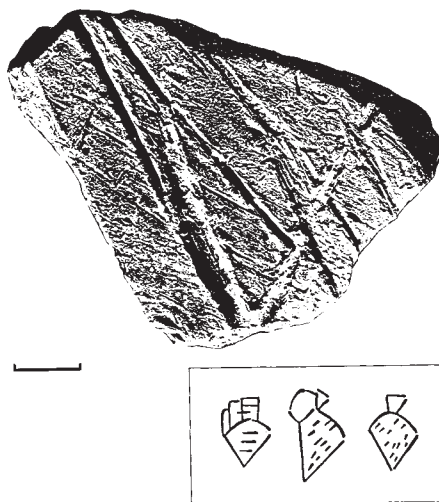
Chemical evidence for ancient beer

SIR — Lower Mesopotamia, comprising the wide alluvial plain of the lower Tigris and Euphrates rivers, was home to one of the oldest literate civilizations in the world — that of the early Sumerian city-states — dating back to the Late Uruk Period (late fourth millennium BC)¹. Irrigation agriculture of domesticated cereals included barley, from which beer is made. Beer was the preferred fermented beverage of the ancient Sumerians². Godin Tepe, a site (excavated by a team from the Royal Ontario Museum) in the nearby Zagros mountains of Iran which has strong Lower Mesopotamian influences, has yielded the earliest chemical evidence for beer. We have discovered a characteristic organic residue inside a pottery vessel which was evidently used for beer fermentation or storage.

During the late fourth millennium BC, the early Sumerians exploited surrounding areas for precious commodities³. At Godin Tepe, the discovery of various artefacts in well-constructed buildings contemporary with the Late Uruk lowland city-states⁴ indicates the presence of early Sumerians (V.R.B., manuscript in preparation).

The people at Godin Tepe, with their close ties to lowland Mesopotamia, were probably also beer drinkers. Carbonized six-row barley was very common at the site⁵ and could have been converted to

beer there. We have observed curious criss-cross grooves covering the interior below the shoulder of a Late Uruk jar from Godin Tepe (see figure). If the jar served as a beer vessel, the grooves may have been designed to retain beer sediments. The early Sumerian sign for beer (*kaš*; ref. 6) shows a jar with similar linear markings (see figure). Together,



Comparison of interior of beer jar from Godin Tepe (Royal Ontario Museum) (top) and Sumerian graphic signs for beer (*kaš*) with respective grooves and internal markings. Scale bar, 1 cm. (Photo courtesy of W. Pratt, Royal Ontario Museum.)

the archaeological and textual evidence strongly suggest that this vessel was a beer container.

Some of the grooves contained a pale yellowish residue which gave a positive test⁷ for oxalate ion. Calcium oxalate, which is soluble in water only with difficulty (6 mg l⁻¹ at 18 °C), is a principal component of 'beerstone' and settles out on the surfaces of fermentation and storage tanks of barley beer⁸. Similar test results were obtained from scrapings from an Egyptian New Kingdom jar (clearly intended for beer according to tomb paintings and reliefs), beerstone from a modern brewer's vat, and pure calcium oxalate.

Spinach and rhubarb, which grow in the Iranian highlands today and presumably grew there in the past, contain substantial amounts of oxalates (5–10%). Compared to cereals, however, they are a minor part of the human diet and we can think of no reason why they should have been stored or processed in the ancient grooved vessel. Although minor amounts of oxalates are widely distributed in nature, exterior scrapings of our Uruk jar and the interiors of vessels from the same site which originally contained a grape product all gave negative results.

The chemical evidence for the earliest beer at Godin Tepe complements the finding of the earliest grape wine there, also dating to the last half of the fourth millennium BC⁹. Analysis of storage and processing containers from earlier periods and from a wider geographical area are needed to determine the ultimate origins and significance of beer production.

Rudolph H. Michel

Patrick E. McGovern

Museum Applied Science Center for
Archaeology (MASCA),
University Museum of Archaeology and
Anthropology,
University of Pennsylvania,
Philadelphia,
Pennsylvania 19104, USA

Virginia R. Badler

Department of Near Eastern Studies,
University of Toronto,
Toronto, Ontario,
Canada M5S 1A1

1. Nissen, H. J. *The Early History of the Ancient Near East, 9000–2000 B.C.* (University of Chicago, 1988).
2. Röhl, W. *Das Bier im alten Mesopotamien* (Gesellschaft für die Geschichte und Bibliographie des Brauwesens, Berlin, 1970).
3. Algaze, G. *Curr. Anthropol.* **30**, 571–608 (1989); **31**, 66–69 (1990).
4. Weiss, H. & Young, T. C. Jr *Iran* **13**, 1–17 (1975).
5. Miller, N. F. in *Progress in Old World Palaeoethnobotany* (eds Van Zeist, W., Wasylikow, K. & Behre, E.-E.) 133–160 (Balkema, Rotterdam, 1991).
6. Green, M. W. & Nissen, H. J. *Zeichenliste der archaischen Texte aus Uruk* (Mann, Berlin, 1987).
7. Feigl, F. *Spot Tests in Organic Applications* (Elsevier, New York, 1954).
8. DeClerck, J. A *Textbook of Brewing* (translator Barton-Wright, K.) (Chapman and Hall, London, 1957).
9. Badler, V. R., McGovern, P. E. & Michel, R. H. in *MASCA Res. Pap. Sci. Archaeol.* **7**, 25–36 (1990).

Universal biology

Richard Dawkins

Artificial Life: The Quest for a New Creation. By Steven Levy. *Cape / Pantheon*: 1992. Pp. 390. £16.99, \$25.

I AM enthusiastic about the embryonic field of artificial life and want to be positive about Steven Levy's book. So let me get the minor carping out of the way first.

"...scientist Harold Thimbleby, a computer scientist at Stirling University..." The adjectival noun is ugly as well as superfluous but journalist Steven Levy, a journalist, is apparently obliged by union rules to shove it in, together with "evolutionary biologist William Hamilton", "Nobel-laureate Niko Tinbergen" and countless others.

The adjectival noun is just a trivial diagnostic of journalese. But Levy also tests positive for that more serious malady of science journalists, Dark Ages syndrome. There is a new wave of heroic young scientists (the ones the journalist has been interviewing and hanging out with). Before they burst on the scene we were mired in the Dark Ages. Left over from the Dark Ages are the Old Guard, traditional scientists with attitudes as hard as their arteries, who control the university jobs and the grant money. Our story is one of doughty deeds by the New Heroes, in the teeth of uncomprehending hostility from the Old Guard.

Occasionally, to be sure, scientific revolutions really do work like this; and in the case of artificial life its luminaries come trailing genuinely swashbuckling pasts: Doyne (pronounced Doan) Farmer and Norman Packard, who typed with their toes on ultra-miniature computers hidden in their shoes in order to break the bank at Las Vegas; Stephen Wolfram, who contemptuously forsook Eton and Oxford and created *Mathematica*, one of the most admired computer programs ever written; Danny Hillis, another boy-wonder, who invented the legendary Connection Machine; and the *éminence rouge* of 'Artificial Life', Christopher Langton himself, who broke all four limbs and most other bits too as an exhibition hang-glider. But the Dark Ages/Young Buccaneer formula has become a lazy cliché among science journalists, and we need to see less of it.

Levy doesn't fall into this cliché quite as heavily as some of his colleagues whose books range from anti-evolution to palaeontology to chaos theory. But he too cannot resist the temptation:

Traditional computer scientists had no trouble ticking off reasons why this could not possibly work; they cited mathemat-

ical principles that theoretically limited the speed gain of parallel processing.

This is a set-up, of course, for a New Hero to come bounding on-stage and confound the establishment, and Hillis duly does. But it is an unsatisfying way of writing, because a mathematical principle is a mathematical principle and we are left wondering how even the Old Guard managed to get it so wrong.

Genetic algorithms could generate robust programs and artificial adaptive phenomena by utilizing the power of evolution. Yet the lords of computer science were slow to bestow their blessings on it...

Who are these 'lords' and do the rest of us really so obsequiously grant them the power to 'bestow blessings'?

In the past, the prejudice against mathematical modeling in theoretical biology might have been justified.

What prejudice against mathematical modelling? What does Levy think theoretical biology mostly is, if not mathematical modelling? See any back issues of *Journal of Theoretical Biology*.

Enough of griping, let us turn this into an interesting point. There is a prejudice lurking in theoretical biology, but it is

not against mathematical modelling. It is more pernicious than Levy realizes, and artificial life goes right to its heart. It is a prejudice against the very idea of a truly theoretical biology, against a biology that strays far from the narrow path of 'data'. I first recognized this prejudice in 1982, at the Darwin Centenary Conference in Cambridge. I presented a paper called "Universal Darwinism", in which I argued that darwinian natural selection is not just the principle that *happens* to underlie life on this planet; natural selection is a *necessary* feature of life anywhere. If organized, apparently purposeful complexity is found anywhere in the Universe, I suggested, some form of Darwinism will be the ultimately responsible force.

In the discussion my lecture was vehemently attacked, not by any Old Guardee but by a distinguished biologist and statistician. It wasn't that he disagreed with the thesis of Universal Darwinism; if he had, that would have been interesting. No, his complaint was that I had not backed up my claim with data! You might as well attack Pythagoras because he didn't sally out into the field and measure a statistical sample of right-angled triangles. In print after the conference, my paper was assailed on similar grounds as "philosophical" — a sufficiently pejorative epithet in some scientific circles. I didn't realize it then because the phrase had not been coined, but my little essay on Universal Darwinism was grounded in the same philosophy as the now aspirant field of artificial life. It is because of this philosophy that we should take the field



Plant programming — P. Prusinkiewicz simulated this *Mycelis muralis* using an L-system program, which adds self-similar new growth with each iteration of the formula. The picture is taken from the colourfully illustrated *Fractals: The Patterns of Chaos* by John Briggs (Simon and Schuster, \$20 (pbk)).

seriously, even if we are not — yet — impressed by any of the attempts to synthesize life in computer or test tube.

Most of biology is, rightly, the study of the way living things actually are. Mammals have twice as many heart chambers as fish. Haemoglobin is red. The list of such facts is too large to be memorized by any individual. You can treat them as just plain true. But could there be a subset of facts that not only are true but *have* to be true (or are, *a priori*, very likely to be true)? If so, isn't there a sense in which this subset is more important than the facts that just happen to be true? My Universal Darwinism conjecture is a candidate. There are others. Nerve spikes convey information by their frequency rather than their amplitude. Is this just another fact or could it be predicted for the functional equivalents of nervous systems on other planets? Hereditary information is digital not analogue and it is transmitted, irreversibly, with exceedingly high fidelity (low mutation rate). Are these just facts about life on Earth, or must they be true of life everywhere? What about the genetic code itself. Is it a frozen accident? Could it have been different? Can we predict that lives on other planets will have detailed similarities such as triplet codes? Sex? A separation between germ line and soma and an equivalent of Francis Crick's 'central dogma'?

It is difficult to answer questions about life in general, because we have a sample of only one to look at: the DNA-coded, protein-engineered, carbon-based life that has evolved over the past three to four billion years on this planet. If there are other lives in the Universe they are probably too far away in space and time for us ever to know about them. The only practical hope we have of studying a sample of lives greater than one is to synthesize them ourselves and let them evolve. Our present technology is such that the digital computer, perhaps especially in parallel, network form, is our best bet as an incubator for artificial life. Some simulations might be realized as robots actively moving, 'feeding' and conceivably reproducing in the real world. One might also hope to use real, wet glassware to develop alternative biochemistries and alternative genetics. All these, and their evolution, belong in the field of artificial life.

Levy's book is an adequate history of the field both before and after it was named. He fittingly begins with John von Neumann's self-reproducing machines and moves, via John Conway and cellular automata, to 1987 when Christopher Langton, realizing that there was a field waiting to be named, convened the first artificial-life conference in Los Alamos. I had the honour to

be invited to this memorable gathering, and I confirm that Levy captures something of its unique, spacy, *Dr Who*-like atmosphere. He does not neglect the serious deliberations of the conference, which were soberly recorded under Langton's excellent editorship in *Artificial Life* (Addison-Wesley, 1989). Levy makes some attempt at analysing the theoretical and philosophical issues, with loads of that 'bottom-up top-down' stuff (the phrases seem instantly to strike a deep chord with everybody except me); but here I think I would rather wait for Langton's own book, due to be published in 1993, probably with the same title again.

You may read *Artificial Life* to learn about gee-whizz achievements so far, but I think you will be disappointed if you do. Give that aspect another few years to develop. A better way for a biologist to approach the field is as an aid to separating contingent facts about life from necessary, or at least highly probable, facts about life. It can help us to understand life in general, and this will illuminate Earth's life in passing. Artificial life emancipates biology from the tyranny of the particular and parochial. □

Richard Dawkins is a fellow of New College, Oxford OX1 3BN, UK.

Revealing ancient language

Nicholas J. Saunders

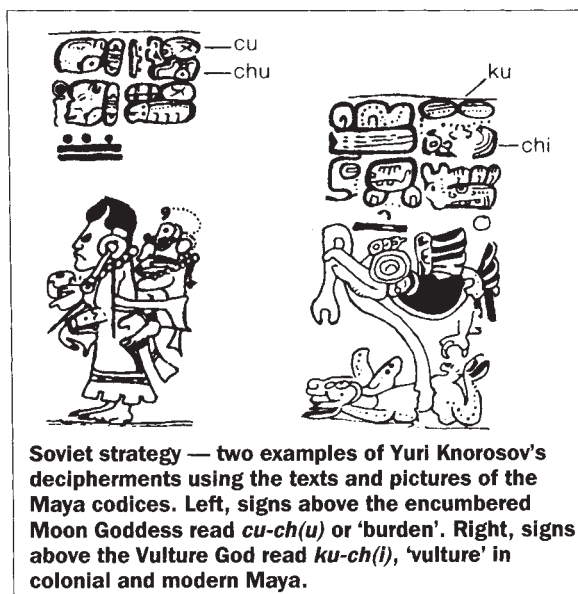
Breaking the Maya Code. By Michael D. Coe. *Thames and Hudson: 1992.* Pp. 304. £14.95, \$29.95.

FEW scholarly achievements are as glamorous as the deciphering of an ancient script, as Jean François Champollion's decoding of Egyptian hieroglyphics and Michael Ventris's cracking of Linear B have illustrated. These triumphs set precedents for a linguistic approach to decipherment and yet were, by turns, misapplied, ignored and eventually redefined by those seeking the grail of decipherment for the Maya hieroglyphics — pre-Columbian America's only complete script. In *Breaking the Maya Code*, Michael Coe tells this compelling story through the colourful personalities of the protagonists.

Despite the relationship between ancient scripts and modern languages established by Champollion, most Maya scholars pointedly ignored surviving Maya tongues as a possible key to decipherment. This despite the occasional prescient insight, such as Constantine Samuel Rafinesque's 1832 letter to Champollion, in which he evidently regards the inscriptions as representing the still extant Maya language. No comprehensive linguistic assault on the Maya script materialized however, and worse setbacks were to come. Coe elegantly describes the developing conflict between those who espoused a linguistic historical approach and others who saw Maya hieroglyphs as unrelated to language

and used solely for calendrical and astronomical purposes. In admirable detail we are presented with an array of mathematically orientated results achieved by such meticulous scholars as Ernst Förstemann, who discovered that the Maya used a vigesimal (base 20) counting system, possessed a 'long count' of years and could account for the movements of Venus. In a similar vein was the discovery by John Teeple in 1925 that the Maya time for an average lunar month was only 33 seconds off the actual value.

The cumulative effect of such discoveries was to stifle occasional dissenting voices such as Cyrus Thomas, who regarded the Maya script as a mixture of phonetic signs and logograms (single symbols representing an entire morpheme, word or phrase). It also gave rise to a distinctive yet wholly untenable view of Maya society. By adhering to the calendrical-astronomical theory, such formidable Mayanists as Sylvanus Mor-



Soviet strategy — two examples of Yuri Knorosov's decipherments using the texts and pictures of the Maya codices. Left, signs above the encumbered Moon Goddess read *cu-ch(u)* or 'burden'. Right, signs above the Vulture God read *ku-ch(i)*, 'vulture' in colonial and modern Maya.

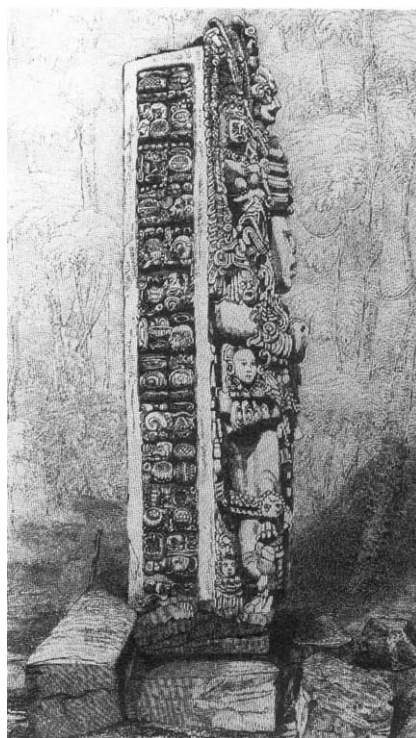
ley and J. Eric S. Thompson were caught in a conceptual straitjacket, believing that Maya society was a peaceful theocracy ruled by a benevolent priesthood. This intellectual élite's apparent lack of interest in recording real or idealized history in favour of esoteric and mythological events made Maya civilization suspiciously unique.

Caught between admiration for Thompson's many scholarly achievements and a dislike of his often pedantic style, Coe nevertheless blames this forceful personality for almost single-handedly blocking the decipherment for half a century. We see how those who dared to espouse a linguistic interpretation were buried under Thompson's biting dismissive rejoinders and vast erudition. Consequently, by 1950 not one named Maya ruler or city was known, and the mass of non-calendrical inscriptions were no better understood than they had been 50 years earlier.

But in 1952, Coe's hero and Thompson's most formidable adversary appeared in the form of the Russian academician Yuri Knorosov. He published convincing evidence that the principles of the Maya script were the same as for the other hieroglyphic systems. Knorosov showed that individual hieroglyphs could at times be phonetic and at other times represent a morpheme (the smallest unit of meaning); that phonetic signs could lessen textual ambiguity; and that glyphs could be inverted for calligraphical reasons. Thompson's uncharitable reaction was to dismiss the breakthrough as nothing less than a Marxist hoax. As Coe perceptively remarks, however, the depth of Thompson's antagonism may well have been due to his suspicion that Knorosov was right.

What Knorosov accomplished for the nature of the Maya script, Tatiana Proskouriakoff achieved for its content. In 1960 she showed that monumental inscriptions recorded history, not astronomy or religion, and that figures graven onto stelae were men and women not gods. By cutting the Gordian knot of Maya epigraphy, Proskouriakoff revealed a decidedly violent world of dynastic rivalry, royal marriages and an astronomically controlled complex of sacrifice and ritual bloodletting.

Not unexpectedly, Thompson's death in 1975 opened the floodgate to new advances. At a series of meetings in the Maya city of Palenque, brainstorming sessions of a new generation of Mayanists resulted in their surmising that if inscriptions reflected Maya language then they should also display the same syntactical structure. Following this line of argument, Linda Schele and Peter Mathews formulated the dynastic history of Palenque by reconstructing the life stories of six successive rulers. There



Cryptic stones — 'Stela A' in Copán, Honduras, as drawn by F. Catherwood, the artist on the expeditions of the American lawyer J. L. Stephens (1805–1852); it was Stephens who first brought the Maya civilization to world attention.

followed an encounter between Schele and a child prodigy named David Stuart, who in eight hours read an inscription that had taken his mentor and others some five years to decipher.

By the mid-1980s, despite disbelieving Mexican scholars and antagonistic North American archaeologists, progress continued apace. The Maya script was finally revealed as being logographic, largely phonetic in content, characterized by the principles of polyvalence (when a single sign has multiple sound values and when a sound is symbolized by more than one sign) and homophony (when multiple signs have the same sound value), and evidently written in a form of the Maya Cholan language family. These breakthroughs also revealed unsuspected details of Maya life, such as the royal status of Maya scribes, a penchant for naming ceramics and temples alike, and the belief in an individual's animal spirit-familiar. The wheel had turned full circle and the Maya were once more speaking for themselves.

Coe has written that rare book — a masterpiece that transcends the boundaries between academic and popular writing. It is a compulsive account of one of the greatest intellectual adventures of the twentieth century. □

Nicholas J. Saunders is in the Department of History, University of the West Indies, St Augustine, Trinidad, West Indies.

Network visions

Terrence Sejnowski

Neural Networks for Perception. Vol. 1: Human and Machine Perception. Vol. 2: Computation, Learning and Architectures. Edited by Harry Wechsler. Academic: 1991. Pp. 520/363. \$59.95, £39.50/\$49.95, £33.

ON the morning of 6 October 1991, a neural network drove a van 35 km along Route 28 north of Pittsburgh, setting a distance record for an autonomous land vehicle. The 'eye' of the network was a camera providing a 30×32-pixel view of the road ahead, with the output controlling the steering. The network had previously been trained for 5 minutes by a human driving along the road; then the network took control and proceeded to steer the van along new roadway at 90 km per hour (the maximum speed of the van). Behind all of this was Dean Pomerleau, then a graduate student in the Department of Computer Science at Carnegie-Mellon University, who used a network with only four hidden units between the visual input and the steering output.

Imagine sometime in the future writing a cheque at a supermarket counter; in the time taken to sweep it through an opening, the bank and account codes are read off the cheque. This is accomplished by a single analog VLSI (very large scale integration) chip, about the size of a thumbnail, which consists of a primitive 'retina' that detects the characters flying by, a 'shifter' that aligns the characters, and a 'cortex' that categorizes them. This development is about to become a reality as Synaptics, a small company in San Jose, California, founded by Federico Faggin and Carver Mead, starts to market its first commercial product. Neural networks are under intense investigation as an engineering approach to difficult problems in vision and as a scientific tool for exploring questions about biological vision. *Neural Networks for Perception* is a two-volume sampler with 37 chapters on recent research in this area.

Consider visual textures. B. Julesz, who pioneered the study of texture discrimination 30 years ago, and D. Williams have a chapter on the ability of humans to segregate textured areas on the basis of differences in the texture elements, such as 'T' and 'L'. J. Malik and P. Perona contribute a chapter showing how a network of nonlinear filters similar to the broadly tuned simple and complex cells found in the early stages of processing in the visual cortex

can account for many of these psychophysical observations. Their insight is that visual cues in early vision (brightness, motion, stereo disparity, colour, texture perception) are based on a redundant multi-resolution, multi-orientation representation of the image, rather than on discrete tokens such as textons, brightness edges or zero crossings as suggested by Julesz, D. Marr and others. But, as pointed out by Williams and Julesz, these models do not adequately account for the perceptual asymmetries observed with some texture pairs when figure and ground are reversed. At higher levels of visual processing, such as form recognition, similar principles may apply. A few global features extracted from a large training database, such as those learned by Pomerleau's driving network, outperform much larger systems using hand-crafted local features.

Several chapters consider the problem of how to develop a large-scale perceptual system from special-purpose heterogeneous networks. Oja explores the use of unsupervised, self-organizing maps to extract relevant features from the input stream. The general problem of how to cascade unsupervised learning in a sequence of visual maps remains an open question. Finkel, Reeke and Edelman study occlusion. They show that a distributed, interactive system, designed along the lines of the visual cortex, can segment images such as the Kanizsa triangle, in which the contours are defined subjectively. Ballard and Whitehead examine simple sequential visual behaviours, such as block manipulation. They explore reinforcement learning, a biologically plausible model of learning based on predicted reward rather than external supervision, and come to the surprising conclusion that reconstructing a detailed internal model of the physical world, an explicit goal in computer vision, might not be desirable. Despite our impression that we possess such an internal model, recent psychophysical experiments suggest that it is an illusion.

What new have we learned about visual perception? Some of the chapters in this collection were driven by traditional problems in computer vision, perhaps with a new twist; others, such as the chapter by Ballard and Whitehead, have considerably altered fundamental assumptions. Vision is more than computational optics because actions affect sensations — Ballard calls this 'animate' vision. In the simplest feedforward models, such as Pomerleau's driving network, the link between action and vision is reflexive, but in more complex systems with feedback loops the link is dynamic, with motor control and short- and long-term memory representations becoming central issues.

Brains live in the real world and perform in real time; neural networks as computational models based on principles from biological systems face the same challenges. For neural networks to survive they need to evolve in two ways. First, most of the research has relied on simulations, which are flexible but inefficient. A hardware infrastructure must develop that is cheaper and more powerful than that based on conventional computational models. Several chapters on VLSI chips for neural networks and optical methods for interconnecting them present promising new technology, but none of these devices is close to industrial strength. Second, new network architectures must be found that scale well with increasing problem complexity. A single network can drive a van along divided highways, but many different networks are required for the range of road conditions encountered every day; the problem of how to coordinate these networks for reliable, efficient control is unsolved. Here we have much to learn from nature, which has produced a diversity of brain structures and adaptive behaviours that we are just beginning to understand. Read *Neural Networks for Perception* for a glimpse of a field in transition, not for a coherent new approach to visual perception. □

Terrence Sejnowski is in the Computational Neurobiology Laboratory, Salk Institute, 10010 North Torrey Pines Road, La Jolla, California 92037, USA.

Visual computation

Computational modelling of vision is also the subject of several other books published in recent months. A few are listed below:

- *Computational Models of Visual Processing* edited by M. S. Landy and J. A. Movshon. MIT Press, \$66.
- *Neural Networks and Image Processing* edited by G. A. Carpenter and S. Grossberg. MIT Press, \$55 (pbk).
- *Computer Vision: A Unified, Biologically-Inspired Approach* by Ian Overington. Elsevier, DFL 185, \$105.50.
- *Understanding Vision: An Interdisciplinary Perspective* edited by Glyn W. Humphreys. Blackwell, £40 (hbk), £16.95 (pbk).

Network overviews

Two introductory books to neural networks and their applications that have recently been published are:

- *Neural Networks: Current Applications* edited by P. G. Lisboa. Chapman and Hall, £29.95 (pbk).
- *An Introduction to the Modeling of Neural Networks* by P. Peretto. Cambridge University Press, £60 (hbk), £24.95 (pbk).

Web of intrigue

Frances Balkwill

The Cytokine Handbook. Edited by Angus Thomson. Academic: 1991. Pp. 425. \$115, £59.

THOSE of us enmeshed in the complex web of the cytokine network may find it difficult to appreciate the problems that this rapidly expanding area must present to the novice. A good guide book is certainly necessary. *The Cytokine Handbook* provides a comprehensive introduction to these cell-modulating proteins, with contents that are of interest not only to the uninitiated but also to those already hooked. If the book was a straight catalogue of the characteristics and actions of individual cytokines it would be of limited use, and would date even more quickly. But here such descriptions have much greater relevance, sandwiched as they are between chapters that place cytokines in the context of the immune system and its evolution, that review the molecular biology of most known cytokines, and that describe some of the therapeutic potential of both cytokines and their receptors.

Any new book on cytokines will quickly date. Important areas such as the soluble cytokine receptors and the cytokine receptor families, and the newer cytokines such as interleukins 8–12, would each merit a chapter had the book been written in 1992 instead of 1990. Curiously, the review copy claims to be a second 1992 printing of a 1991 publication, although it is not clear whether the reprinting is due to great demand or to the publisher's being overly cautious about the initial size of the print run. An update of the text would certainly be warranted before a third printing.

Over the past two years, several new cytokine books have been published, all of them excellent general guides, but shorter than *The Cytokine Handbook*. There is a tendency for these guides to contain a few key references at the end of each chapter rather than to reference individual points. This can be frustrating for those who want to delve into the original literature. By contrast, each chapter of *The Cytokine Handbook* refers to well over a hundred original papers, making the book as a whole an important reference source. It provides a useful guide to cytokines and their receptors, and, provided the contents are regularly updated, could well become the definitive introductory text. □

Frances Balkwill is in the Biological Therapies Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK.

Is the Antarctic ice sheet growing?

S. S. Jacobs

A common public perception is that global warming will accelerate the melting of polar ice sheets, causing sea level to rise. A common scientific position is that the volume of grounded Antarctic ice is slowly growing, and will damp future sea-level rise. At present, studies supporting recent shrinkage or growth depend on limited measurements that are subject to high temporal and regional variability, and it is too early to say how the Antarctic ice sheet will behave in a warmer world.

WHAT will happen to the polar ice sheets in the face of increasing global temperatures? Higher temperatures will cause more ice to melt, with a concomitant global sea-level rise. At the same time, moderate temperature increases could increase precipitation at high latitudes, so that more water is 'locked up' as snow on the polar caps and sea level falls accordingly. If this latter idea seems at odds with the well-known relationship between sea level and continental ice-sheet size over geological time, it nonetheless has a factual basis. Accumulation rates in Antarctica are known to vary in proportion to mean annual air temperature above the surface inversion layer¹, consistent with lower rates during the last glacial maximum². Antarctica's total annual water budget is several times that of Greenland, and the snow that falls on its grounded ice is equivalent to $\sim 5 \text{ mm yr}^{-1}$ of global sea-level change. Observed sea-level rise (SLR) has been in the $1.0\text{--}2.4 \text{ mm yr}^{-1}$ range over the past century, with a 'best guess' of 1.5 mm yr^{-1} (ref. 3). At the lower limit most SLR can be accounted for by ocean thermal expansion and the melting of temperate and Greenland margin glaciers³. This would imply that the Antarctic ice sheet has recently been near equilibrium, but larger SLR estimates³⁻⁵ suggest a significant unexplained deficit. If our future climate is moderately warmer and wetter than at present, increased precipitation on Antarctica, where average temperatures would remain below freezing year-round, may lock up more water and retard SLR^{3,6-8}.

There have been several recent reports of polar ice-sheet growth, based on satellite altimeter measurements over Greenland⁹, mass budget analyses¹⁰ and positively correlated increases

in air temperature and accumulation¹¹. These reports must be interpreted with some caution. Altimeter observations are of limited duration and are most accurate for gently sloping accumulation regimes within the equilibrium line. The continental applicability of measured increases in accumulation on Antarctica, which exceed those expected from any associated warming¹¹, may be compromised by high spatial and temporal variability and by changing distance from moisture sources. Mass budgets also vary and are based on sparse measurements over limited time periods. Indeed, it can plausibly be demonstrated from available accumulation, calving and melting data that the Antarctic ice sheet has been slowly shrinking over the past few decades¹². Growth of this ice sheet during the next century will depend on its own internal dynamics and on changes in the Southern Hemisphere which affect the atmosphere and ocean circulations, the sea-ice cover and ice-shelf stability. Here I present a brief review of recent observations of surface accumulation on Antarctica. I conclude that as yet it is too early to say with confidence whether the ice sheet has recently been growing or shrinking, given the variability in accumulation patterns and the larger uncertainties in melting and calving.

Accumulation and temperature changes

Analyses of four ice cores from the Wilkes Land sector of East Antarctica show that recent accumulation rates (1975-85) are $\sim 20\%$ above the long-term mean (1930-85)¹¹. Those ice cores were taken from a limited region within 300 km of the coastline, but are consistent with earlier results from the same general

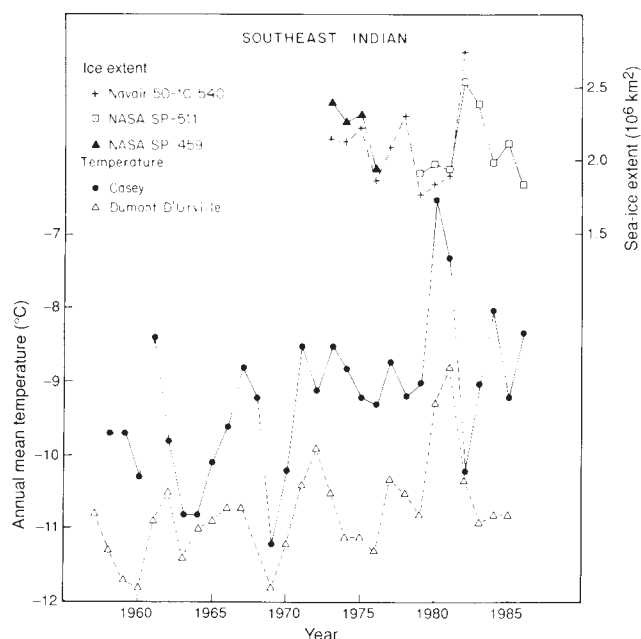


FIG. 1 Monthly average sea-ice extent at the seasonal maximum in the Southeast Indian sector of the Southern Ocean (90° – 160° E)^{18,55,56} and annual mean temperature at the coastal Casey and Dumont d'Urville stations⁵⁷ (Fig. 2). Interannual temperature changes are regionally coherent, larger toward the west and negatively correlated with the Fig. 2 sea ice anomalies during the higher variability period after 1979. Statistical analyses for the 1973-87 period in this region show that above-normal air temperatures can be related to below-normal ice coverage, and that lag correlations exist between seasonal anomalies¹⁶.

area of East Antarctica¹³. The two longest cores extended into ice deposited before 1850 and displayed a large variability on timescales of up to 100 years¹¹. Although only 18 km apart, they also showed widely different accumulation rates of 570 and 1,160 kg m⁻² yr⁻¹ from 1975–85. Such high temporal and spatial variability, which can be expected near the coastline, makes calculations of accumulation changes fairly sensitive to the choice of sampling site and reference interval.

All four ice cores revealed low accumulation during the 1955–65 period, and a subsequent increase paralleling a rise in mean Antarctic air temperatures¹¹. Warmer air can carry more moisture, and precipitation tends to be positively correlated with temperature over Antarctica⁶. But temperatures are sometimes dominated by cyclonic activity, and in some regions accumulation is greatest during the coldest months^{14,15}. This may apply to Casey Station, the only Antarctic location that has experienced a significant rising temperature trend in all seasons¹⁶. As Casey, at 66° 15' S, 110° 32' E, is nearest the East Antarctic ice cores, accumulation increases at the drill sites may be greater than at most locations around Antarctica. There is also evidence that regional temperature trends are out of phase, and that the

Antarctic-wide temperature trend is biased by the preponderance of coastal records¹⁷.

Accumulation and sea-ice variability

The East Antarctic accumulation increases are too large to be accounted for solely by rising air temperature, leading to the suggestion that a primary cause may be an intensification of cyclonic activity in the circumpolar low-pressure trough near Casey Station^{11,15}. Increased cyclonic activity would probably cause divergence in the sea-ice field, altering its area or extent, enhancing the rise of warmer deep water and the transfer of moisture into the atmosphere. Sea-ice changes should be detectable in satellite passive microwave records that have been available since 1973, long enough to look for regional patterns. Over this period, seasonal maximum sea-ice extents near the Wilkes Land coast, a sector where precipitation is greatest during winter¹⁵, show slight declines punctuated by a large increase in 1982 (Fig. 1). Warmer (colder) temperatures at the two coastal Wilkes Land stations do tend to be associated with less (more) sea ice. During the warmer years at Casey, ice extents were generally less than 2.0×10^6 km². The coldest year in this interval (-10.2°C in 1982) coincided with a maximum ice extent of 2.5×10^6 km², 38% greater than the mean of the four warmest years. This suggests that sea-ice cover may also have been greater during the 1957–72 period (Fig. 1), when the average temperature at Casey (-9.8°C) was close to the colder 1982 value.

These sea-ice and air-temperature relationships in the Southeast Indian sector are consistent with other reports¹⁶ and broadly correspond with the increase in Wilkes Land accumulation since the late 1960s. The relative proximity of the Wilkes Land coast to the mean northern edge of the sea ice may foster a closer atmospheric link between the open sea and continental precipitation in this region than, for example, in the Weddell or Ross Seas (Fig. 2). In addition, lowered ice concentrations within a sea-ice field will bring moisture sources much closer to the continent than the northern ice limit or mean annual distance to the open ocean¹⁹. An extreme example of this and its possible impact on the oxygen isotope and accumulation records is afforded by the Weddell Polynya of the mid-1970s²⁰. Coastal polynyas formed by offshore winds could also be a factor if the warmed and moistened air that has passed over them is incorporated into cyclonic storms that track onto the continent.

Decreased sea ice in one Southern Ocean sector is commonly balanced by an increase elsewhere¹⁸. This is illustrated by the alternating zonal anomalies in the August 1981 panel of Fig. 2. These interannual changes in the sea-ice cover may contribute to spatial and temporal variability in accumulation on the Antarctic continent. Numerous glacial ice cores near the coastline could thus become useful indicators of regional changes in sea ice cover if calibrated against the post-1972 satellite records. Decreases in circumpolar sea-ice extent reported since 1973 have usually been short-term or too small to be significant^{16,21,22}. Nonetheless, positive trends in air temperature have been associated with a movement of the northern ice limit to higher latitudes²³. Future warming coupled with large reductions in sea-ice extent could lead to large increases in Antarctic accumulation⁸. Minor warming in the southern atmosphere coupled with sea-ice growth⁷ could have the opposite effect.

Accumulation on the Antarctic ice sheet

From the foregoing, it is not apparent that recent accumulation increases in East Antarctica^{11,13} will be representative of the entire continent. Those data are consistent with measurements on the Antarctic peninsula²⁰, the western part of which is also near the sea-ice edge and has recently experienced one of the largest positive temperature anomalies in the Southern Hemisphere²⁴. But there is another side to the question, with some regions having experienced accumulation rate decreases^{25,26} or little apparent change^{27,28} over the past several decades

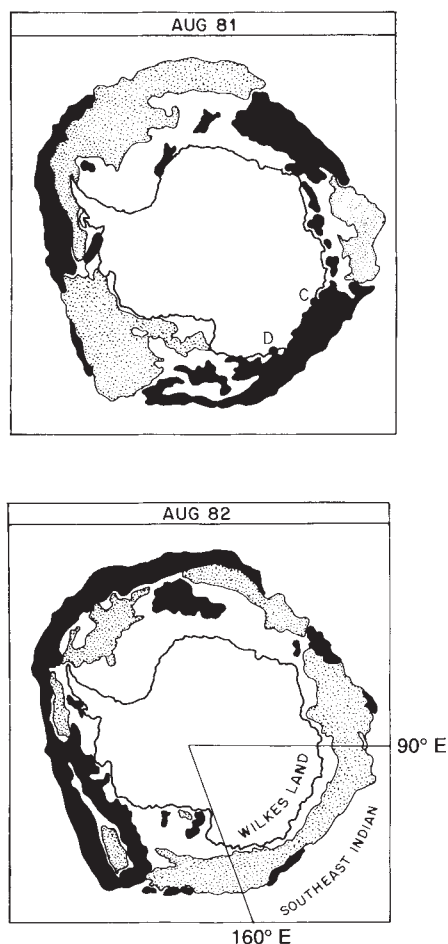


FIG. 2 Mean Antarctic sea-ice concentration anomalies during a late winter month in 1981 and 1982 relative to the 1979–87 August average⁵⁶. The heavy (light) shading in the Southeast Indian sector of the 1981 (1982) panel indicates sea ice concentrations 5–60% lower (higher) than average. Adjacent months show similar patterns, with generally low ice cover in this region during 1980 and 1981 when surface air temperatures at coastal stations C and D were warm, but ~25% above average in 1982 (ref. 56) when temperatures were coldest (Fig. 1). Note the relative proximity of the northern ice limit to the coastline in this sector, and the alternating circumpolar bands of higher and lower than average ice concentrations, particularly in August 1981.

(Fig. 3). Locations with accumulation increases seem to outnumber those with decreases, but this sampling may not be representative, and there tend to be more sites near the coastal regions. Other records also reveal a low accumulation period in the mid-1950s¹⁴, from which it might be inferred that much of the recent increase represents a return to 'normal'. Alternatively, it might be an example of the 30% variations in accumulation rate that can be observed for decades²⁹.

To understand possible continent-wide changes, it is desirable to have an accepted baseline or best estimate for annual accumulation during some specific period. The choice of such a value for net accumulation is not straightforward, given the variability in both deposition and sampling. This can be seen from the range of estimates cited in Table 9.6 of the IPCC report³. Those surface balance values range from 1,468 to 2,168 gigatonnes (Gt) yr⁻¹ and are represented as accumulation only on the grounded ice. Most of the estimates are derived from a Scott Polar Research Institute (SPRI) compilation. A review of the five original sources (A to E) suggests additional anomalies, and a possible resolution.

(A) A surface balance of 2,088 Gt yr⁻¹ is cited versus "about 2×10^3 km yr⁻¹ [sic]... with a discrepancy most likely in the range of 0 to +20%" in the original source³⁰, where "balance velocities" are calculated from the SPRI data. Total influx is not specified as applicable only to the grounded ice, but this can perhaps be inferred from a later report⁸.

(B) A surface balance of 2,168 Gt yr⁻¹ is quoted. This value is hard to find in the original³¹, but is close to the figure of 2,189 Gt yr⁻¹ in ref. 32. The author describes the behaviour of the Antarctic ice sheet "and its ice shelves" over the last glacial-interglacial period, by means of a model run initiated 160 kyr ago. He indicates difficulty in parameterizing the accumulation rate, necessitating "interpolation and smoothing of original measurements (1,000 data points from SPRI)".

(C) A surface balance of 2,158 Gt yr⁻¹ is given for the grounded ice, and the source³³ shows that a substantially larger 2,633 Gt yr⁻¹ is applicable to the calving front. A 1987 addendum

to this report includes the caveat that the relatively large values may be due to an "upward bias" and that "a definite update of the original steady-state description of the ice sheet therefore remains a future task".

(D) A surface balance of 1,468 Gt yr⁻¹ ($\pm 10\%$) seems to be very low, in part because it excludes the Antarctic Peninsula³. In addition, the authors³⁴ selected between incompatible data sets that should not be averaged, and accounted for areas of deflation and net ablation in the coastal zone not shown on small-scale maps³⁵ (M. Giovinetto, personal communication). Adding to this value accumulation on the Peninsula and ice shelves^{12,36} gives 2,144 Gt yr⁻¹.

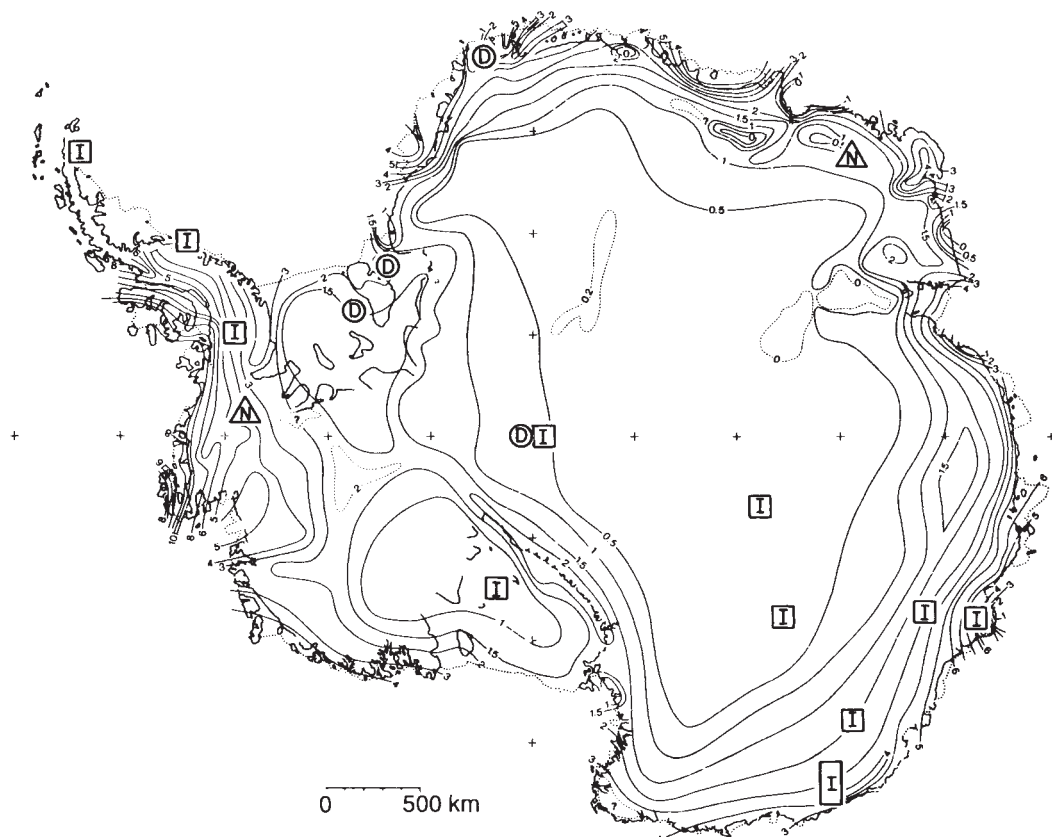
(E) A surface balance of 1,817 Gt yr⁻¹ appears consistent with units of km³ in the source text³⁷, if not with kg³ [sic] in its Table V. The authors apportioned 86% of total precipitation on the grounded ice, as opposed to 75% in (D), and found that only in the interior could surface mass balance be parameterized reliably by a linear multiple-regression analysis. Current mass supply would be 2,104 Gt yr⁻¹ with the ice shelves included.

It seems that source B may represent total ice-sheet accumulation, close to values given by D and E when the latter are adjusted to include the Antarctic Peninsula and ice shelves. If source A is applicable only to the grounded ice, that would make it consistent with C and both would be substantially higher than the other estimates. Arguably the best choice is the one that initially appeared the least consistent (D), but was within 8% of the mean of 12 accumulation studies performed from 1971–85 (ref. 35). If this value is used with the adjustments noted above, grounded and total ice-sheet accumulations become 1,528 and 2,144 Gt yr⁻¹ (ref. 12). The former is at the lower limit of the 30% accuracy assigned to a 2,200 Gt yr⁻¹ value in the IPCC report for accumulation on the grounded ice (Table 9.3 in ref. 3).

Mass balance of the Antarctic ice sheet

To assess the mass balance of the entire ice sheet, contemporary attrition terms must be compared with net accumulation. One

FIG. 3 Surface mass balance rates³⁴, in units of $100 \text{ kg m}^{-2} \text{ yr}^{-1}$, and locations of ice cores from which recent accumulation changes have been reported. 'I' indicates an accumulation increase^{11,13,20}, 'D' a decrease^{25,26} and 'N' little or no apparent change^{27,28}. Specific values are not cited because of the variety of units and time intervals in the original sources. Accumulation is highest within a few hundred kilometres of the Antarctic coastline, where projections of future accumulation increases are also highest⁸, but where ice residence time on the continent will be lowest. Correlations have been reported between sea level pressure and snow accumulation in East Antarctica, possibly related to changes in the frequency and tracks of coast-parallel cyclones⁵⁸.



approach¹⁰ is to integrate the outflow across a defined cross-section or exit 'gate' in each drainage basin, through which the ice velocity must be known. Velocities vary with position and time, however, and mass transfers can occur upstream of the gates, which are often not located at the grounding lines. A recent result with this technique showed an excess mass input of 40–400 Gt yr⁻¹ over the grounded ice sheet¹⁰. An alternate method, also subject to substantial uncertainty, is to sum all wastage products—iceberg calving, basal melting minus freezing, and runoff. One of these attrition terms often considered to be minor, basal melting of the ice shelves, is now known to exceed 500 Gt yr⁻¹, which moves it into the major category. The melting rate and a calving estimate derived from satellite and census data, in combination with the accumulation of 2,144 Gt yr⁻¹, shows a deficit of ~470 Gt yr⁻¹ for the current balance of the entire ice sheet¹².

The usual caveats apply this negative estimate of mass balance. Inferences must be drawn from limited observations, individual budget components can have error bars from 20 to 50%, and existing records are not yet long enough for significant trends to be separated from the large natural variability. Nonetheless, it is important to define baseline best estimates for all budget terms, particularly the basal meltwaters whose flow went for a long time unnoticed far beneath the surface of the sea. The rate of *in situ* melting of the ice shelves, the temporal variability of which is unknown, may be sensitive to thermohaline changes in the ocean^{38–40}. In the future, comprehensive elevation data from satellite-borne laser altimeters should allow the ice-sheet volume to be monitored. But if the flow is temporally variable over short periods and the ice is thickening in some areas and thinning in others⁴¹, a long time series may be required to extract a meaningful signal for the entire continent.

The Antarctic ice sheet and sea-level change

The IPCC report³ concludes that it is unknown whether the Antarctic ice sheet is currently in balance and whether it has contributed to sea-level rise over the past century. These are separate but related issues, because only the grounded parts of an ice sheet can directly and significantly influence sea level by gaining or losing mass. Suppose that the grounded ice is in fact growing at a rate of 40–430 Gt yr⁻¹ (refs 10 and 11, assuming 1 mm of sea-level change is roughly equivalent to 360 Gt of water), and that the entire ice sheet is shrinking at a rate of 470 Gt yr⁻¹ (ref. 12). The floating ice shelves and ice tongues would then be losing mass at a rate of 500–900 Gt yr⁻¹. The historical record⁴² and large calving events over the past decade¹² might be cited as evidence for this¹⁰, but other lengthy ice fronts have also been advancing recently without major calving^{43,44}. In addition, that attrition rate would be roughly equivalent to the disintegration of the Wordie ice shelf⁴⁵ twice each year. There is limited evidence that some ice-front positions have varied little over the past 150 years⁴³ and that others have receded since the last glaciation⁴⁶. It has also been proposed that rapid fluctuations in the size of the Ross ice shelf could occur without significantly affecting the inland ice sheet⁴⁷. But if the ice shelves are now waning and can effectively lever outflow across the grounding line^{38,48}, that would be cold comfort for the more distant future when increasing ice-sheet flow could become a principal factor in sea-level change⁸.

Linking recent Antarctic ice-sheet accumulation increases to a lowering of sea level^{10,11} highlights the problem of accounting for observed global sea-level rise over the past century. That rise has been in the range of 1 to 2.4 mm yr⁻¹, with higher values more applicable to the later decades⁴⁹. A 'best estimate' of 1.5 mm yr⁻¹ sea-level rise over the past century included 1.05 mm yr⁻¹ for ocean thermal expansion and the meltwater from temperate glaciers and Greenland⁵, but left 0.45 mm yr⁻¹ of rise unexplained. An additional negative 0.1–1.2 mm yr⁻¹ from Antarctica^{10,11} would increase that unknown contribution to 0.55–1.65 mm yr⁻¹. This would imply that thermal expansion

has been greatly underestimated or that 200–600 Gt yr⁻¹ of water have been leaking into the ocean from unidentified sources. That could place a heavy burden on sources such as groundwater withdrawal, wetland loss or desertification, not generally cited as significant contributors⁵⁰. Alternatively, if the ice shelves had a negative balance of ~300 Gt yr⁻¹ (half their surface balance rate), accumulation on the grounded ice were 1,500–1,700 Gt yr⁻¹ (refs 10, 12) and calving were 2,000–2,200 Gt yr⁻¹ (refs 3, 12), the current grounded ice budget would still be sufficiently negative to account for the 0.45 mm yr⁻¹ anomaly in sea-level rise.

It is certainly possible that the Antarctic ice sheet has grown as sea level has risen over the past century, but the opposite conclusion seems equally tenable. Likewise, increased accumulation and air temperatures in East Antarctica and on the Antarctic peninsula could be the expected sign of a negative feedback that could slow the rise in the sea level. This is a common element of general circulation models^{8,51} and a factor in ice sheet growth for moderate warming³². But general circulation model results that project more Southern Hemisphere warming penetrating into the deep ocean⁵² could imply more ice attrition and less of an accumulation increase in the future. If increased subsurface ocean temperatures are accompanied by increased basal melt rates, ice shelves could thin and decrease the back-pressure on the grounded ice^{53,54}. Alterations in the atmospheric forcing and Southern Ocean thermohaline structure are likely to change the ocean circulation as well as the sea-ice distributions. The nature of those changes and their probable influence on the mass balance of the Antarctic ice sheet remain to be determined. □

S. S. Jacobs is at the Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA.

1. Robin, G. *Phil. Trans. R. Soc. (Lond.)* **B280**, 143–168 (1977).
2. Lorus, C. et al. *Nature* **316**, 591–596 (1985).
3. Warrick, R. & Oerlemans, J. in *Climate Change. The IPCC Scientific Assessment* (eds Houghton, J. T. et al.) (Cambridge Univ. Press, 1990).
4. Peltier, W. R. & Tushingham, A. M. *J. geophys. Res.* **96**, 6779–6796 (1991).
5. Douglas, B. C. *J. geophys. Res.* **96**, 6981–6992 (1991).
6. Robin, G. in *The Greenhouse Effect, Climatic Change and Ecosystems* (eds Bolin, B. et al.) (SCOPE 29, ICSU, WMO, Wiley 1987).
7. Manabe, S., Stouffer, R. J., Spelman, M. J. & Bryan, K. *J. Clim.* **4**, 785–818 (1991).
8. Budd, W. F. & Simmonds, I. in *Proc. int. Conf. Role of Polar Regions in Global Change, June 11–15, 1990* (eds Weller, G. et al.) 489–494 (Univ. of Alaska, 1991).
9. Zwally, H. J., Brenner, A. C., Major, J. A., Bindshadler, R. A. & Marsh, J. G. *Science* **246**, 1587–1589 (1989).
10. Bentley, C. R. & Giovinetto, M. B. in *Proc. int. Conf. Role of Polar Regions in Global Change, June 11–15, 1990* (eds Weller, G. et al.) 481–488 (Univ. of Alaska 1991).
11. Morgan, V. I., Goodwin, I. D., Etheridge, D. M. & Woakey, C. W. *Nature* **354**, 58–60 (1991).
12. Jacobs, S. S., Hellmer, H. H., Doake, C. W. S., Jenkins, A. & Frolich, R. *J. Glaciol.* **38** (in the press).
13. Pourchet, M., Pinglot, F. & Lorus, C. *J. geophys. Res.* **88**, 6013–6020 (1983).
14. Bryazgin, N. N. in *Climate of Antarctica* (ed. Dolgin, I. M.) (Russ Transl. Ser. **41**, Balkema, Rotterdam, 1986).
15. Bromwich, D. *Rev. Geophys.* **26**, 149–168 (1988).
16. Weatherly, J. W., Walsh, J. E. & Zwally, H. J. *J. geophys. Res.* **96**, 15119–15130 (1991).
17. Mosley-Thompson, E. in *Climate since AD 1500* (eds Bradley, R. S. & Jones, P. D.) (Routledge, London 1992).
18. Zwally, H. J. et al. *NASA Spec. Publ.* 459 (1983).
19. Giovinetto, M. B., Waters, N. M. & Bentley, C. R. *J. geophys. Res.* **95**, 3517–3531 (1990).
20. Peel, D. A. in *Climate since AD 1500* (eds Bradley, R. S. & Jones, P. D.) (Routledge, London 1992).
21. Gloersen, P. & Campbell, W. J. *Nature* **352**, 33–36 (1991).
22. Zwally, H. J., Parkinson, C. L. & Comiso, J. C. *Science* **220**, 1005–1012 (1983).
23. Jacka, T. H. & Budd, W. F. in *Proc. int. Conf. Role of Polar Regions in Global Change, June 11–15, 1990* (eds Weller, G. et al.) 63–70 (Univ. of Alaska, 1991).
24. Folland, C. K. et al. in *Climate Change 1992. The Supplementary Report to the IPCC Scientific Assessment* (eds Houghton, J. T. et al.) (Cambridge Univ. Press, 1992).
25. Graf, W., Reinwarth, O., Oerter, H. & Dyurgerov, M. in *Filchner-Ronne Ice Shelf Programme 4*, (ed. Miller, H.) (Alfred Wegener Inst., Bremerhaven, 1990).
26. Jouzel, J., Merlivat, L., Petit, J. R. & Lorus, C. *J. geophys. Res.* **88**, 2693–2703 (1983).
27. Mosley-Thompson, E. et al. *J. Glaciol.* **37**, 11–22 (1991).
28. Satow, K. & Watanabe, O. *Ann. Glaciol.* **14**, 256–260 (1990).
29. Lambert, G., Petit, J. R., Pourchet, M., Sanak, J. & Lorus, C. in *Int. Conf. Evolution of Planetary Atmospheres & Climatology of the Earth, Nice 1978* (French with Engl. summary) (Centre Nat'l d'Etudes Spatiales, Toulouse, 1979).
30. Budd, W. F. & Smith, I. N. in *Glaciers, Ice Sheets and Sea Level: Effects of a CO₂-induced Climatic Change* (Nat. Acad. Press, Washington DC, 1985).
31. Huybrechts, P. *Ann. Glaciol.* **14**, 115–119 (1990).
32. Huybrechts, P. & Oerlemans, J. *J. Clim. Dynam.* **5**, 93–102 (1990).
33. Radok, U., Brown, T. J., Jensen, D., Smith, I. N. & Budd, W. F. *On the Surging Potential of Polar Ice Streams. Part IV. Antarctic Ice Accumulation Basins and their Main Discharge Regions* (Coop. Inst. Res. Envir. Sci., Univ. of Colorado, Boulder 1986, 1987).
34. Giovinetto, M. B. & Bentley, C. R. *Antarct. J. U.S.* **20**, 6–13 (1985).
35. Giovinetto, M. B., Bentley, C. R. & Bull, C. C. *Antarct. J. U.S.* **24**, 7–13 (1989).

36. Frolich, R. M. in *The Contribution of the Antarctic Peninsula Ice Sheet to Sea Level Rise*, Commission of the European Communities Rep. (ed. Morris, E.) (British Antarctic Survey, in the press).
37. Fortuin, J. P. F. & Oerlemans, J. *Ann. Glaciol.* **14**, 78–84 (1990).
38. Thomas, R. H. in *Glaciers, Ice Sheets and Sea Level: Effects of a CO₂-induced Climatic Change* (Nat. Acad. Press, Washington DC, 1985).
39. Hellmer, H. H. & Oelbers, D. *J. Antarct. Sci.* **1**, 325–336 (1989).
40. Jenkins, A. *J. geophys. Res.* **96**, 20671–20677 (1991).
41. Alley, R. B. & Whillans, I. M. *Science* **254**, 959–963 (1991).
42. Zakharov, V. G. *Polar Geogr. Geol.* **12**, 297–311 (1988).
43. Jacobs, S. S., MacAyeal, D. R. & Ardai, J. L. Jr. *J. Glaciol.* **32**, 464–474 (1986).
44. Lange, M. A. & Kohnen, H. *Ann. Glaciol.* **6**, 187–191 (1985).
45. Doake, C. S. M. & Vaughan, D. G. *Nature* **350**, 328–330 (1991).
46. Kellogg, T. B. & Kellogg, D. E. in *Int. Assoc. Hydrol. Sci. Publ.* **170** (eds Waddington, E. D. & Walder, J. S.) (1987).
47. Jezek, K. C. *J. geophys. Res.* **89**, 409–416 (1984).
48. Bindshadler, R. A. et al. *West Antarctic Ice Sheet Initiative, Science and Implementation Plan*, NASA Conf. Publ. 3115 (NASA, Washington DC, 1991).
49. Barnett, T. P. in *Sea-Level Change*, Studies in Geophysics (Nat. Acad. Press, Washington DC, 1990).
50. Meier, M. *Nature* **343**, 115–116 (1990).
51. Mitchell, J. F. B., Manabe, S., Meleshko, V. & Tokioka, T. in *Climate Change: The IPCC Scientific Assessment* (eds Houghton, J. T. et al.) (Cambridge Univ. Press, 1990).
52. Stouffer, R., Manabe, S. & Bryan, K. *Nature* **342**, 660–662 (1989).
53. Thomas, R. H. *J. Glaciol.* **24**, 273–286 (1979).
54. Lingle, C. S., Schilling, D. H., Fastook, J. L., Paterson, W. S. B. & Brown, T. J. *J. geophys. Res.* **96**, 6849–6871 (1991).
55. Naval Oceanography Command Detachment, *Sea Ice Climatic Atlas: 1*, Antarctic, Navair 50-1C-540 (1985).
56. Gloersen, P. et al. *Arctic and Antarctic Sea Ice, 1978–1987: Satellite Passive Microwave Observations and Analysis* (NASA SP-511, in the press).
57. Jones, P. D. & Limbert, D. W. S. **TR038** (U.S. Dept of Energy, Washington DC, 1987).
58. Enomoto, H. *Clim. Change* **18**, 67–87 (1991).

ACKNOWLEDGEMENTS. I thank NASA for financial support, B. Batchelder for drafting, and Z. Feng, M. Giovinetto, V. Gornitz, H. Hellmer, A. Jenkins, G. Kukla, V. Morgan and G. Robin for reviews.

ARTICLES

Stabilization of sea urchin flagellar microtubules by histone H1

Luc Multigner*, Jean Gagnon†, Alain Van Dorsselaer‡ & Didier Job*

*Département de Biologie Moléculaire et Structurale, Laboratoire du Cytosquelette, Unité INSERM 244, Centre d'Etudes Nucléaires de Grenoble, BP 85 X, 38041 Grenoble Cedex, France

†Institut de Biologie Structurale, 41 avenue des Martyrs, 38027 Grenoble Cedex 1, France

‡Laboratoire de Spectrométrie de Masse Bio-Organique, Faculté de Chimie, 1 rue Blaise Pascal, 67008 Strasbourg Cedex, France

Complex microtubule assemblies are essential components of eukaryotic cilia and flagella. They are extremely stable and are not affected by agents that normally induce polymer disassembly. The molecular basis of this microtubular stability is unknown, and it is not related to any feature of the constitutive tubulin. In sea urchin sperm flagella, axonemal microtubules are found to be stabilized by a protein identical to histone H1, a result that defines a new role for this histone and provides evidence for a concerted evolution of chromatin and microtubular structures.

A FUNDAMENTAL property of microtubules is that, depending on the type of structure in which they participate, they can either be labile, with an assembly state that is quite sensitive to subtle variations in their environment, or very stable^{1,2}. This difference in stability is central to their function^{3,4}. The complex microtubule assemblies in centrioles or in axonemes of cilia and flagella are formed from stable polymers^{1,2}, but cytoplasmic and mitotic microtubules are generally labile, showing a dynamic instability³. Drugs that promote microtubule disassembly or stabilization have a profound influence on cell organization and cell division². The molecular basis of axonemal microtubule stability is unknown. It is retained by isolated microtubule fragments and does not require the integrity of the complex axonemal architecture⁵. Stable microtubules contain covalently modified tubulin, detyrosinated at the α -subunit carboxy terminus^{6,7} and acetylated at the lysine residue at position 40 of the same subunit⁸, but these modifications are the consequence and not the cause of microtubule stability^{9,10}. When axonemal microtubules are detyrosinated¹¹ and acetylated¹², their constitutive tubulin reassembles *in vitro* into labile microtubules which disassemble at low temperature¹³. We therefore investigated whether stability of axonemal microtubules could be induced by some specialized proteins, as in the case of brain microtubules stabilized by STOP protein¹⁴.

Before any such stabilizing molecules can be characterized, it is necessary to solubilize the axonemal microtubule assemblies as their constitutive subunits using chaotropic agents and/or ionic detergents. Proteins solubilized in this way cannot directly be assayed for their ability to stabilize microtubules, so we have used a filter assay^{14,15} to characterize a single protein of M_r 34,000 (34K) isolated from sea urchin sperm flagellar axonemes. Structural and immunological data reveal that this 34K protein

is identical to nuclear histone H1. We show that the 34K protein is an axonemal component and not a contaminating nuclear protein. Addition of the 34K protein or histone H1 to microtubules assembled *in vitro* from pure brain tubulin induces as high a degree of polymer stability as is found in axonemal microtubules. These results indicate that histone H1 has an unexpected and important role in axonemal microtubule stabilization.

Microtubule stabilizing activity in axonemes

We used a variety of fractionation procedures to study sperm flagellar axonemes from the sea urchin *Paracentrotus lividus* (Fig. 1). Membrane-free axonemes were resuspended in a high-ionic-strength buffer (0.6 M KCl) to solubilize the outer dynein arms and the central pair of microtubules^{13,16}, which we confirmed by negative-stain electron microscopy (data not shown). Thermal denaturation of the KCl-extracted axonemes solubilized some inner dynein arms and most of the B tubules of outer doublets¹⁷. The remaining material was primarily A-tubules and was fractionated using solutions containing Sarkosyl detergent and increasing concentrations of urea^{18,19}. Solubilized proteins were separated at each step from the insoluble material by centrifugation and microtubule-stabilizing activity was assayed in the supernatants.

Most microtubule-stabilizing activity appeared in two major protein fractions: a 0.5% Sarkosyl extract (lane 3, Fig. 1b; fraction I) and in a 0.5% Sarkosyl extract containing 5 M urea (lane 7, Fig. 1b; fraction II). We observed interspecies variations in activity when using this extraction procedure: fraction I was the more active from the sea urchin *Arbacia lixula*, for example (data not shown). The activity in fraction I from *P. lividus* was further purified using gel filtration and reversed-phase

high-pressure liquid chromatography (RP-HPLC) (Fig. 2a and b) to give a protein of M_r 34K (34K-I) (see Fig. 2b insert); a protein band of the same apparent molecular weight in fraction II (lane 7, Fig. 1a) was further purified by RP-HPLC (Fig. 2c) (34K-II protein; see Fig. 2c insert). Other axonemal protein fractionation procedures using different ionic detergents always yielded a microtubule stabilizing activity that copurified with a 34K protein (data not shown). One Sarkosyl extract containing 2 M urea (lane 6, Fig. 1a) contained a set of polypeptides which, based on their apparent M_r and solubility, correspond to tektins¹⁹; this fraction was inactive in our assay.

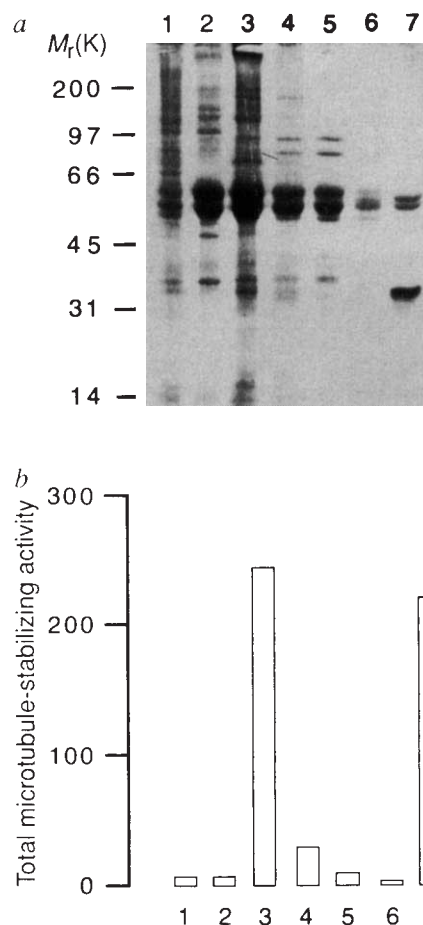
A rabbit polyclonal antibody was raised against the 34K-II protein. Immunoblot analysis showed that the antibody reacted with the 34K-I protein, suggesting that the 34K-I and 34K-II proteins are related (lanes 1 and 2, Fig. 2d). The same antibody

reacted with a single 34K protein band on immunoblots of total axonemal proteins separated by SDS-PAGE (lane 3, Fig. 2d).

Axonemal 34K proteins and histone H1 compared

We determined the N-terminal sequence of the 34K-I protein (Fig. 3b) and found that the N terminus of 34K-II protein was blocked. Therefore we cleaved the 34K-II protein by reaction with CNBr and sequenced the fragments (Fig. 3b). The sequences were checked against the NBRF (National Biomedical Research Foundation) data bank and we found that they could all be aligned with histone H1 from sperm of the sea urchin *Parechinus angulosus* and that there was complete identity between the N-terminal sequence of the 34K-I protein and histone H1. The CNBr peptides derived from the 34K-II protein

FIG. 1 Fractionation of sea urchin sperm axonemal proteins (a), and determination of their microtubule-stabilizing activity (b). a, Coomassie brilliant blue (CBB)-stained 10% SDS-polyacrylamide gel. The numbers on the left indicate M_r in thousands. Lanes 1 to 7 show the protein composition of each of the axonemal fractions: lane 1, 0.6 M KCl-soluble fraction; lane 2, heat-soluble fraction; lanes 3 and 4, 0.5% Sarkosyl-soluble fractions (two successive extractions); lanes 5 and 6, 0.5% Sarkosyl/2 M urea-soluble fraction (two successive extractions); lane 7, 0.5% Sarkosyl/5 M urea-soluble fraction. b, Total microtubule-stabilizing activity in each fraction defined as the ratio of the total volume of the fraction to the volume of the aliquot yielding a stabilization of 50% of microtubules under standard incubation conditions. **METHODS.** Mediterranean sea urchins *P. lividus* were collected at the Marine Station of Villefranche-sur-Mer, France. Mature male gonads were removed and macerated gently. Extruded sperm was filtered through a single layer of nylon mesh (0.25-mm pore size). All subsequent steps were at 4 °C. The dry semen was diluted in 20 vols filtered (0.22- μ m pore size) sea water and centrifuged at 200g for 15 min to eliminate coarse debris and pigmented phagocytes. Sperm cells were separated from the seminal plasma by centrifugation (at 1,500g for 15 min). Cell pellets were resuspended in one original volume of HMCK buffer (10 mM HEPES, pH 7.0, 4 mM MgCl₂, 1 mM CaCl₂, 0.1 M KCl, 0.02% β -mercaptoethanol) and the membranes removed by addition of 4 vols of the same buffer containing 1% Triton-X100. After a 1-min incubation, the suspension was homogenized using ~10 strokes in a Dounce homogenizer. Sperm heads and unbroken cells were pelleted by centrifugation (3,000g, 10 min). Axonemes were then recovered by centrifugation (15,000g, 10 min) and resuspended in two original volumes of HMCK buffer. These differential centrifugation steps were repeated until the resuspended sperm axonemes were free from sperm heads when observed by dark-field microscopy. The last pellet of axonemes was diluted in a small volume of storage buffer (10 mM HEPES, pH 7.0, 0.1 M KCl, 0.05% β -mercaptoethanol, 1 mM dithiothreitol (DTT), 20% glycerol) and stored at -80 °C at a protein concentration of 10 to 15 mg ml⁻¹. The preparation of axonemes was controlled for head contamination by counting the number of axonemes (using β -tubulin immunofluorescence) and of intact or swollen heads (using Hoescht 33258 as a DNA marker). A typical preparation of flagellar axonemes showed less than one head for 8,000 full-length (40–50 μ m) axonemes. Axonemal proteins were fractionated in the presence of the protease inhibitors 1 mM phenylmethylsulphonyl fluoride, 1 μ g ml⁻¹ leupeptin, 1 μ g ml⁻¹ pepstatin, 1 μ g ml⁻¹ chymostatin and 0.1 μ g ml⁻¹ trypsin inhibitor. Axonemes were pelleted at 15,000g for 10 min and resuspended at a 10 mg ml⁻¹ concentration in 10 mM Tris-HCl, pH 8.0, 0.6 M KCl, 0.1 mM EDTA, 1 mM DTT. Samples were extracted twice for 30 min at 4 °C with occasional stirring, and centrifuged at 100,000g for 30 min. Both supernatants were mixed and dialysed against 100 vol TED buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA, 1 mM DTT) using a membrane dialysis with a 3,500 M_r cut-off. This extract is referred to as KCl-soluble fraction. Pellets were washed in 10 vols TED buffer, resuspended in 1 mM Tris-HCl, pH 8.0, 0.1 mM EDTA, 1 mM DTT, at a protein concentration of 10 mg ml⁻¹, and the B-tubule thermally denatured at 40 °C for 5 min. Samples were chilled on ice and centrifuged at 100,000g for 1 h; the supernatant is referred to as the heat-soluble fraction. The insoluble material was sequentially extracted (30 min at 0 °C with occasional stirring) using 0.5% Sarkosyl and increasing concentrations of urea (0, 2 and 5 M) in TED buffer. At each step, solubilized and insoluble protein fractions were separated by centrifugation (100,000g, 1 h). Pellets were resuspended in a 10-vol excess (v/v) of extraction buffer. In each fraction, microtubule-stabilizing activity (as defined by induction of microtubule resistance to cold-temperature-induced disassembly) was measured¹⁴. Results are expressed as per cent of stability¹⁴.



could be aligned with the conserved hydrophobic central core of the same histone. We already knew that the 34K-I and 34K-II proteins had the same apparent molecular weight, the same behaviour on RP-HPLC, and that they showed immunological cross-reactivity. The sequence analysis showed that these two proteins also shared homology with a sea urchin sperm histone H1. The only apparent difference between the 34K-I and the 34K-II protein was the presence of a blocked N terminus in the case of the 34K-II protein. This can be readily accounted for by the use of urea during the isolation procedure of the 34K-II protein. The conclusion is that the 34K-I and the 34K-II proteins are the same molecule and that they are analogous to histone H1. To confirm this finding, we made use of the solubility of histone H1 in perchloric acid, a property that distinguishes it from the other histones²⁰. We checked that the 34K-I and the 34K-II proteins were soluble in perchloric acid (data not shown), then ran parallel experiments in which isolated sperm nuclei or axonemes from *P. lividus* were extracted with perchloric acid. In both cases this resulted in the isolation of a single 34K protein (lanes 1 and 2 of Fig. 3a) which cross-reacted with the anti-34K-II polyclonal antibody (lanes 3 and 4 of Fig. 3a). We sequenced the amino terminus and some of the CNBr fragments from the nuclear and axonemal perchloric acid-soluble proteins (Fig. 3b).

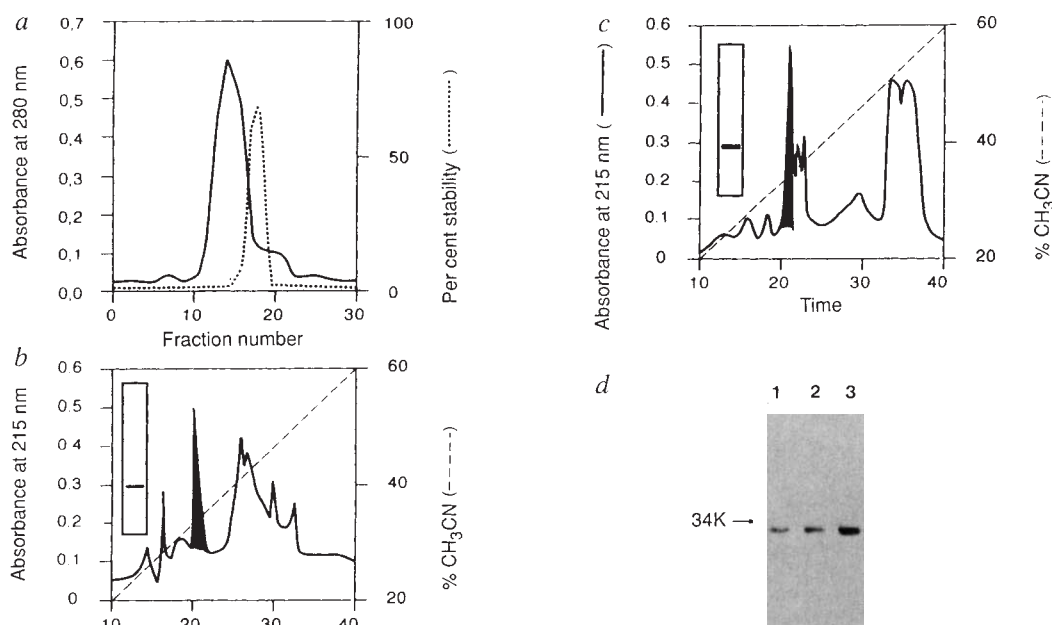
FIG. 2 Purification of axonemal microtubule-stabilizing activities. *a*, Separation of microtubule-stabilizing activities contained in 0.5% Sarkosyl-soluble fractions by gel filtration on Biogel P200. Fractions were assayed for protein content and microtubule-stabilizing activities. *b*, Further purification of the microtubule-stabilizing material by reversed-phase high-performance liquid chromatography (RP-HPLC). Active fractions from the gel filtration column were eluted at 32% acetonitrile (hatched area) and contained a single protein band of M_r 34K (34K-I protein) as determined by 12% SDS-PAGE (insert). *c*, Fractionation of the 0.5% Sarkosyl, 5M urea-soluble fraction by RP-HPLC. Active fractions were eluted at 32% acetonitrile (hatched area) and contained a single protein band of M_r 34K (34K-II protein) as determined by 12% SDS-PAGE (insert). *d*, Immunoblot analysis of purified 34K-I and 34K-II proteins (lanes 1 and 2 respectively), and of total axonemal proteins (lane 3) using a rabbit polyclonal antibody directed against the 34K-II protein.

METHODS. The 0.5% sarkosyl-soluble fractions (lanes 3 and 4, Fig. 1) were pooled. A 1-ml aliquot containing 5 mg of protein was loaded onto a Biogel P200 (Bio-Rad) column (85 cm $\times$ 1.1 cm) equilibrated and developed in TED buffer containing 0.5% Sarkosyl and supplemented with 1 mM PMSF. Flow rate was 1.2 ml per hour and fractions collected were 0.6 ml. Protein elution was monitored by optical density measurement at 280 nm. The microtubule-stabilizing activity was determined in an aliquot of each column fraction. Active fractions were precipitated by cold acidified acetone (-80°C , 10 mM HCl) and centrifuged at 15,000g for 30 min. Pellets were washed twice with cold acetone, dried under a stream of N_2 and resuspended in 6 M guanidine-HCl just before RP-HPLC. The 0.5% Sarkosyl, 5 M urea-soluble fraction was dialysed against 100 vol TED buffer containing 0.5% Sarkosyl, precipitated with cold acidified acetone, and the pellets resuspended in 6 M guanidine-HCl. Typically, 50 μg proteins were loaded into a Baker-Bond (endcapped) C_8 wide-pore analytical column (300 $\text{\AA}$, 4.6 mm $\times$ 250 mm) connected to a Beckman Gold System HPLC. Elution was with a linear gradient of aqueous

As expected, sequence homology was found between both perchloric acid-soluble proteins as well as with histone H1 from *P. angulosus*, from which we concluded that axonemes contain a 34K protein which is homologous to histone H1. But there was a possibility that the axonemal 34K protein and histone H1 sequences diverged in their unsequenced regions, so we analysed both proteins by electrospray mass spectrometry. The profile of the 34K protein showed a single series of multiply charged ions corresponding to M_r 25,488 $\pm$ 12 (Fig. 3c). A sample of histone H1 injected under the same conditions also gave a single series of multiply charged peaks corresponding to an M_r of 25,481 $\pm$ 4 (Fig. 3d). This result ruled out any significant difference in the primary structure of both proteins.

Analysis of axoneme purity

We were concerned that the 34K protein, homologous to histone H1, could have been due to sperm head contamination in the axoneme preparation. In intact sea urchin sperm cells, the ratio of histone H1 to tubulin is $\sim 1/5$ (ref. 21). In purified axonemes the ratio of 34K-I and 34K-II proteins to tubulin is between 2/100 and 10/100 of tubulin, as estimated from the yield after purification and from SDS-PAGE respectively. In our preparations, we determined the head-to-axoneme ratio to be 1/8,000



acetonitrile (broken lines) in constant 0.1% trifluoroacetic acid at a flow rate of 1.0 ml min⁻¹. Protein elution was continuously monitored at 215 nm using a diode array spectrophotometer. Eluted peaks were evaporated to dryness under a stream of N_2 and resuspended in TED buffer. They were assayed for microtubule-stabilizing activity and analysed by 12% SDS-PAGE. To produce a rabbit polyclonal antibody directed against the 34K-II protein, 100 μg HPLC-purified 34K-II protein purified by 10% SDS-PAGE was homogenized in 10 mM Tris-HCl, pH 8.0, and one rabbit was injected with 20 μg 34K-II protein with Freund's complete adjuvant at four subcutaneous sites. It was boosted 4 times at one-week intervals by injections of 20 μg of proteins in Freund's incomplete adjuvant. Total axonemal proteins were solubilized in 1% SDS and boiled in the presence of 5% β -mercaptoethanol. For immunoblot analysis, proteins were separated by 12% SDS-PAGE and electrophoretically transferred onto PVDF Immobilon membranes using a transferring buffer⁴⁰ containing 0.1% SDS. Non-specific protein binding sites were blocked with PBS, 0.1% Tween 20. Membranes were incubated with rabbit anti 34K-II antiserum (1/20) in the presence of heparin (100 μg ml⁻¹). Visualization was achieved by application of an alkaline phosphatase-conjugated goat anti-rabbit IgG secondary antibody. Under these immunoblotting conditions the preimmune sera gave no positive signal.

(see legend to Fig. 1). If the 34K protein in axonemes were a result of head contamination, its ratio to tubulin should be $1/8,000 \times 1/5 = 1/40,000$. This is about three orders of magnitude less than the ratio actually observed (2/100 to 10/100).

Qualitative analysis confirmed that sperm head contamination was not the source of the 34K protein in axoneme preparations (Fig. 4). Although only histone H1 is soluble in perchloric acid, all histones are soluble in 0.2 M sulphuric acid²². We quantified the relative amounts of H1 and core histones and of the 34K protein in sulphuric acid extracts of axonemes, using nuclear sulphuric acid extracts as standards in SDS-PAGE (lane 1, Fig. 4a) and RP-HPLC (Fig. 4b). The amounts of core histones in the nuclear and axonemal extracts, containing comparable amounts of histone H1 and 34K protein respectively, are compared. Histone H1 and 34K protein were estimated from SDS-PAGE (lanes 1 and 2 of Fig. 4a, and lanes 1 and 2 of Fig. 4d), immunoblotting (lanes 3 and 4 of Fig. 4a, and lanes 3 and 4 of Fig. 4d), and absorbance measurements (Fig. 4b and c). If histones were present in axonemal preparations as a result of sperm head contamination, the nuclear and axonemal extracts should have contained identical amounts of core histones, but the set of core histones evident on SDS-PAGE and HPLC profiles from the nuclear extracts was absent in the axonemal

extracts, strongly indicating that the 34K protein in axonemal preparations was not a result of head contamination. This conclusion was confirmed by immunolocalization (see later).

Microtubule stabilization by histone H1

We used filter assays to determine the microtubule stabilizing effect of an increasing concentration of 34K protein or histone H1 on preformed pure tubulin microtubules (Fig. 5a). Results were similar for both proteins and 50% stabilization was obtained at a molar ratio of protein to tubulin of $\sim 1/15$.

Immunofluorescence of pure tubulin microtubules stabilized with either the axonemal 34K protein or histone H1 gave similar images (Fig. 5b-e). These stabilized microtubules were found to resist cold temperature (0°C) and to survive freezing at -80°C, which is a property of axonemal microtubules^{2,5}; they were also resistant to up to 5 mM Ca^{2+} in the presence or absence of calmodulin (up to 10^{-5} M).

In these experiments, either histone H1 or the axonemal 34K protein was added to preformed microtubules. When they were added to unassembled tubulin, turbidity developed but only a few microtubules formed. Microscopy revealed the presence of abundant and mostly amorphous tubulin aggregates (not shown). Thus histone H1 and the axonemal 34K protein can stabilize preformed microtubules but, at least under our

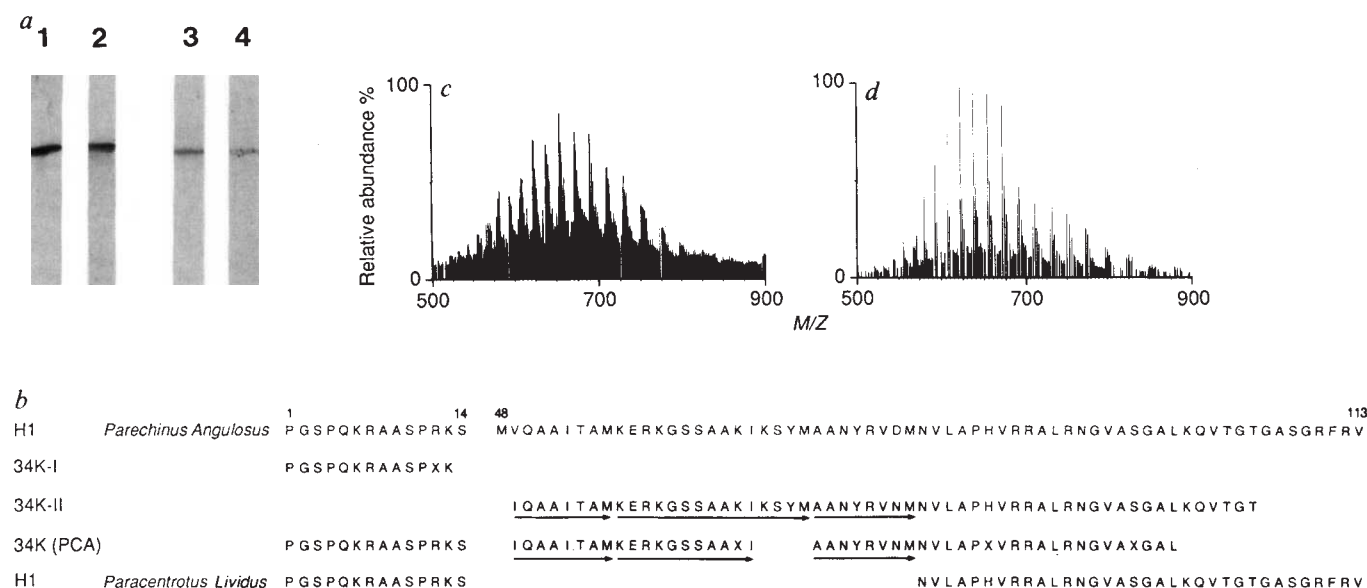


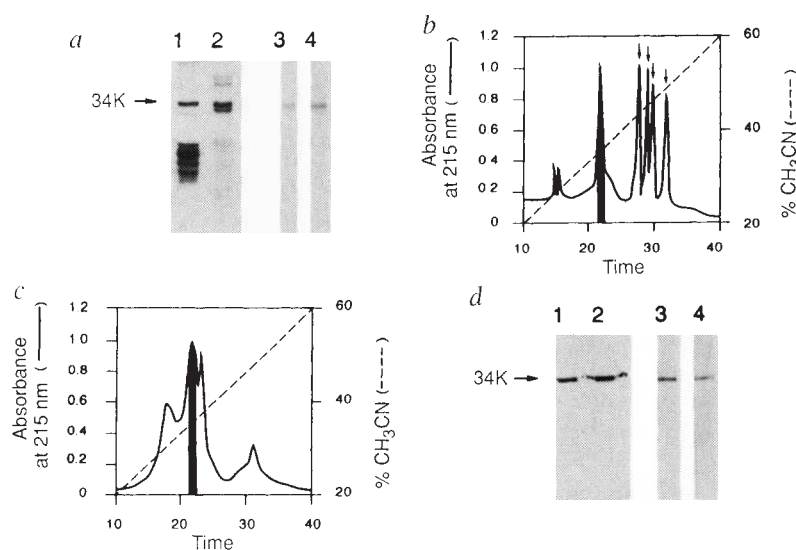
FIG. 3 Comparison of histone H1 with the axonemal microtubule-stabilizing proteins. **a**, 12% SDS-PAGE (lanes 1 and 2) and immunoblot analysis (lanes 3 and 4) of perchloric acid extracts (5% v/v) of sperm nuclei (lanes 1 and 3) and of axonemes (lanes 2 and 4). Immunoblots were probed with rabbit 34K-II antiserum. **b**, Partial amino-acid sequences (single-letter code) of sperm nuclei histone H1 and the axonemal microtubule-stabilizing proteins 34K and axonemal 34K protein, extracted into perchloric acid (PCA). Sequences were aligned on the basis of homology with the N-terminal fragment and the central hydrophobic core (arrows) of the sea urchin *P. Angulosus* histone H1 taken from ref. 41. **c** and **d**, Electrospray mass spectrometry analysis of purified 34K (PCA) protein (**c**) and nuclear histone H1 (**d**). Both spectra are characterized by a high number of charges, which is expected for proteins rich in basic residues.

METHODS. Sperm nuclei were purified from the first pellet of sperm heads obtained as described in Fig. 1 legend. The pellet was resuspended in 2 M sucrose, 10 mM CaCl_2 , 1% Triton X-100 and sheared in an omnimixer for 2 min on ice. After dilution, sperm nuclei were centrifuged through a 1.7 M sucrose cushion for 45 min at 10,000g in a swing rotor. The pelleted nuclei were washed and resuspended at 5×10^9 per ml in a buffer composed of 50 mM Tris-HCl, pH 8.0, 0.5 mM DTT, 1 mM EGTA, 25% glycerol and 1 mM PMSF. The aliquoted nuclei were frozen and stored at -70°C until used. Perchloric acid extraction was done by diluting either purified sperm nuclei

or axonemes with an equal volume of 10% perchloric acid (v/v). Samples were extracted for 4 h at 4°C and centrifuged at 15,000g for 30 min. Supernatants were made 20% in trichloroacetic acid (TCA) with 0.25 vol cold 100% TCA (v/v). After 30 min on ice, precipitated proteins were centrifuged for 15 min at 15,000g. Pellets were washed with cold acidified acetone and twice with acetone and dried under N_2 . Perchloric acid-extracted proteins were purified by RP-HPLC as for Fig. 2 before sequencing and mass spectrum analysis. Proteins in 25% acetic acid were degraded in an automated sequencer Applied Biosystem Model 477 equipped with on-line phenylthiohydantoin derivatives analyser. Aliquots of the 34K-II protein and of the perchloric acid-extracted proteins were dissolved in 70% formic acid and digested with a 100 molar excess of CNBr for 24 h at room temperature under N_2 . Digests were separated by RP-HPLC (data not shown), and individual peptides sequenced. Data bases were searched using DNASTAR. Electrospray was run on a VGB10-Q quadrupole mass spectrometer (BioTech, Manchester, UK) on samples in 1% acetic acid, 50% methanol at a final concentration of $10 \text{ pmol } \mu\text{l}^{-1}$. Injection ($10 \mu\text{l}$) was at a flow rate of $5 \mu\text{l min}^{-1}$ and several scans were averaged to obtain the best signal-to-noise ratio. The extracting cone voltage was set at 90 volts to diminish salts adducts often observed with very basic proteins. Calibration was by separate injection of horse myoglobin. The resolution ($m/\Delta m$) of the instrument was 1,000.

FIG. 4 Core histone content in nuclei and axonemes. *a*, SDS-PAGE of sulphuric acid extracts from sperm nuclei (lane 1) and from axonemes (lane 2). The corresponding immunoblots were analysed using rabbit 34K-II antiserum (lanes 3 and 4). *b*, RP-HPLC of sulphuric acid-soluble extracts from sperm nuclei. The elution positions of histone H1 (shaded area) and core histones (arrows) are indicated. *c*, RP-HPLC of sulphuric acid-soluble axonemal extracts. The major peak (shaded area) corresponding to the elution position of histone H1, whereas no definite peaks were seen at positions corresponding to the nucleosomal core histones. *d*, SDS-PAGE (lanes 1 and 2) and immunoblot analysis (lanes 3 and 4) of the proteins contained in the hatched peaks of RP-HPLC from nuclear (lanes 1 and 3) or axonemal (lanes 2 and 4) sulphuric acid extracts.

METHODS. Purified sperm nuclei (10^9 per ml) or axonemes (5 mg ml^{-1}) were mixed with an equal volume of cold 0.4 M sulphuric acid and stirred for 15 h at 4°C . The acid-insoluble material was removed by centrifugation ($15,000g$, 15 min) and soluble proteins were precipitated by 20% TCA. Precipitates were pelleted, washed with cold acetone and dried. For RP-HPLC, proteins were resuspended in 10 mM HCl . RP-HPLC chromatography, 12% SDS-PAGE gels and immunoblots are described in Fig. 2 legend.



experimental conditions, they inhibit tubulin polymerization by triggering the formation of amorphous tubulin aggregates.

Immunolocalization

The rabbit polyclonal anti-34K-II protein antibody stained the entire length of the sperm flagellum (Fig. 6*a-d*), but there was no staining with preimmune serum (data not shown). Similar staining patterns were observed in other sea urchin species such as *A. lixula* and *Lytechinus pictus* (P. Huitorel and C. Gagnon, personal communication). In the case under study, the antibody showed patterns similar to those with an anti-tubulin antibody when axonemes were splayed, strongly suggesting that the 34K protein was actually present on microtubules. Nuclei did not stain with our anti-34K-II protein antibody (Fig. 6*f*), but it has been found that anti-histone H1 antibodies do not stain sperm nuclei²³, histone H1 being apparently not accessible to anti-

bodies in sperm chromatin. To test the consistency between the immunoblotting and immunofluorescence data, we unravelled histone H1 epitopes in nuclei and found that sperm nuclei stained with our antibody after treatment with sulphuric acid (Fig. 6*h*). Finally, the polyclonal antibody was affinity-purified on a column containing covalently linked nuclear histone H1. The purified antibody stained membrane-free axonemes brightly (Fig. 6*i*), whereas the IgGs in the column flow-through were negative (data not shown). This purified antibody was also tested on sea urchin embryo cilia (Fig. 6*j*), where it distinctly and specifically stained the cilia, suggesting that in sea urchins microtubule stabilization by histone H1 is not restricted to sperm axonemes.

We also used the purified antibody for electron microscopy and immunogold labelling to locate the 34K protein in membrane-free axonemes. Results confirmed the specific association

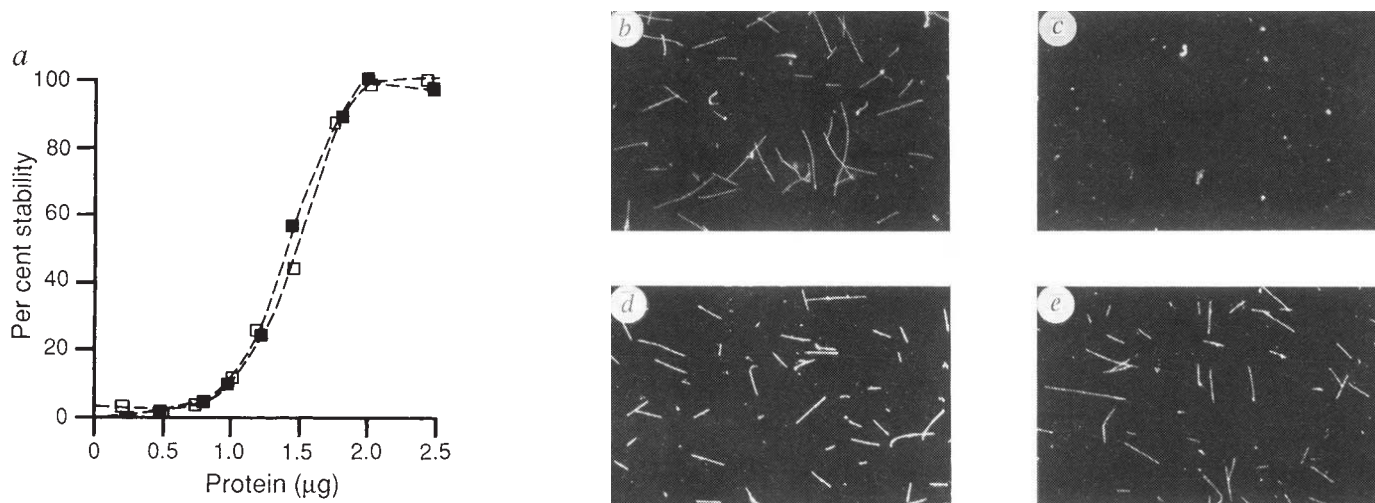


FIG. 5 Pure tubulin microtubule stabilization by the axonemal 34K protein and histone H1. *a*, Assay of microtubule stabilization by increasing concentrations of 34K protein (squares) or of histone H1 (filled squares). *b-e*, Immunofluorescence of microtubules before (*b*) and after (*c-e*) exposure to cold temperature in the absence (*c*) or presence of the axonemal 34K

protein (*d*) or of histone H1 (*e*).

METHODS. Microtubule stabilization was assayed as described in Fig. 1 legend, except that pure tubulin microtubules were assembled as in ref. 15. For immunofluorescence, microtubules were crosslinked and processed as described in ref. 42.

of this protein with axonemal structures but the resolution was not high enough to indicate its position accurately relative to individual microtubules (data not shown).

Other species

To see whether the presence of histone H1 in sperm flagella was confined to sea urchins, we immunostained other cell types using our histone H1 immunoaffinity-purified anti-34K-II protein antibody as a primary antibody. These studies involved two widely divergent ciliates, *Paramecium tetraurelia* and *Euplotes aediculatus*, and a green alga, *Chlamydomonas reinhardtii*. There was a clear positive staining of cilia in *Paramecium* and *Euplotes* and of flagella in *Chlamydomonas* (data not shown). Interestingly, confocal microscopy indicated that no other microtubular structures stained distinctly in optical sections of these cells, with the exception of basal bodies (data not shown), which are structurally related to axonemes. These organelles are easily recognizable in *Paramecium*. Furthermore, a 24K protein, also

corresponding to histone H1, can be extracted from *Chlamydomonas* flagella by perchloric acid (J. L. Rosenbaum and K. A. Johnson, personal communication).

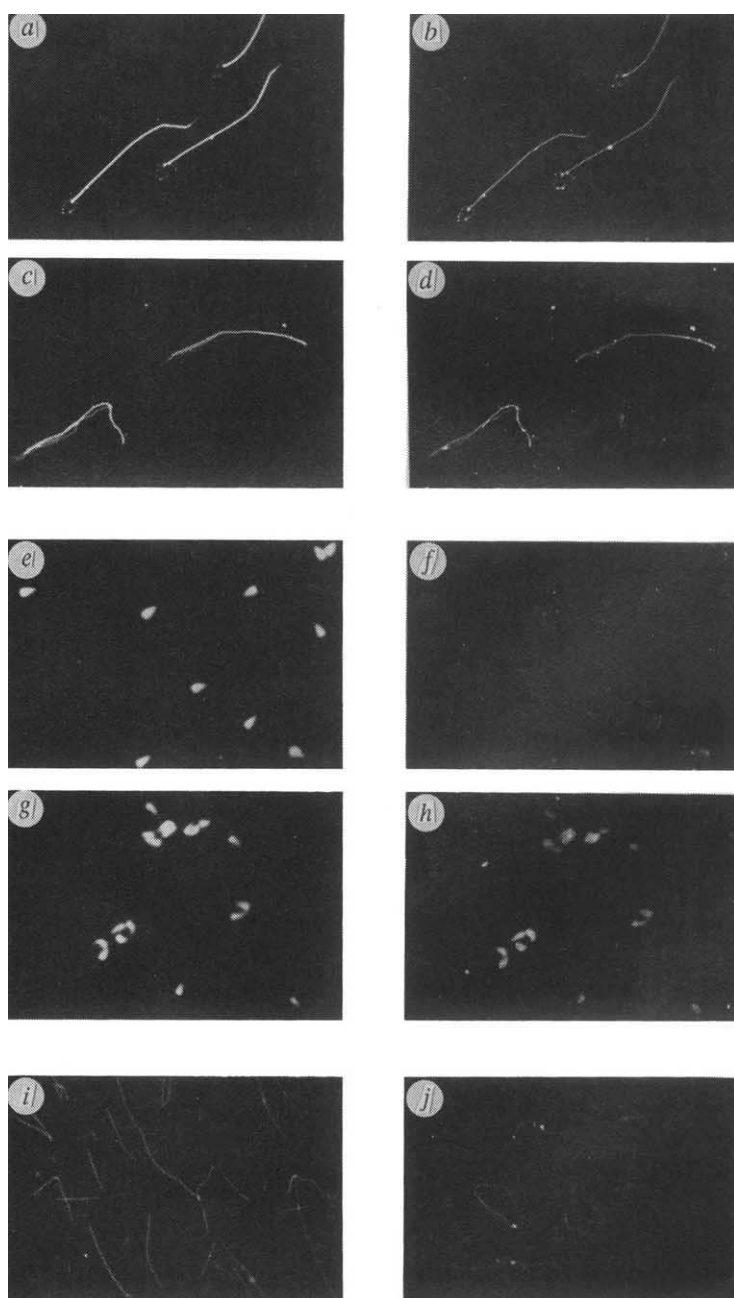
Discussion

In an exhaustive study of microtubule-stabilizing activities contained in sea urchin sperm flagella, we found only one effector which we identified as histone H1. We show that histone H1 is present in relatively large amounts in flagella associated with microtubular structures and conclude that sea urchin axonemal microtubules are stabilized by histone H1. We provide evidence that this applies to axonemal microtubules from other animal cells and to green algae, an observation which we believe has important evolutionary and mechanistic implications.

It is thought that the emergence of microtubular structures was one of the landmarks for the transition between prokaryotes and eukaryotes^{24,25}. Flagella appeared at an early stage of eukaryotic cell evolution, probably at the same time as the

FIG. 6 Immunofluorescence localization of the axonemal 34K protein. Double immunolabelling (a–d) of whole sperm cells showing intact (a and b) or splayed (c and d) axonemes, using an anti-tubulin antibody (a and d) and the anti-34K-II protein antibody (b and c). In control double-staining experiments, in which the anti-34K-II protein antibody was replaced by the preimmune serum from the same rabbit, only the tubulin staining was visible (data not shown). Control single-staining experiments using the 34K-II antibody showed the same staining pattern as in b (data not shown). Double staining (e–h) of purified sperm nuclei before (e and f) and after (g and h) incubation with 0.2 M sulphuric acid, using Hoechst 33258 (e and g) and the 34K-II antibody (f and h). Immunolabelling of isolated sperm axonemes (i) and detached cilia from sea urchin embryos (j) using histone H1 immunoaffinity-purified anti-34K-II protein antibody. Control experiments using the flow-through of the immunoaffinity column gave negative results (data not shown).

METHODS. Sea urchins and sperm cells were obtained as described in the legend to Fig. 1. A suspension of eggs, isolated by dissecting gonads in filtered sea water, were filtered through a 130- μ m nylon mesh and the eggs allowed to settle. Just before insemination, dry semen was resuspended in filtered sea water. Insemination was done at a concentration of 4×10^4 eggs per ml in a tenfold excess of sperm cells under gentle swirling. After hatching, ciliated embryos were recovered by centrifugation (at 500g for 5 min). Treatment of nuclei with sulphuric acid was at 0 °C for 3 h in 0.2 M acid. Nuclei were then diluted in 10 mM HEPES, pH 7.2, 0.1 M KCl and processed for immunofluorescence. Immunoaffinity purification of the anti-34K-II protein antibody was carried out by loading an aliquot of the serum on an Affigel-10 column containing immobilized nuclear histone H1, purified as described in the legend to Fig. 3. Specific antibodies were eluted with 0.1 M glycine, pH 2.5. The purification factor achieved by this procedure was >200-fold. Sperm cells, nuclei and axonemes were sedimented onto 14-mm round coverslips at 12,000g for 15 min and fixed in methanol at –20 °C for 6 min. Sea urchin embryos were similarly processed, except that they were centrifuged at 500g for 10 min. Tubulin was stained using a mouse monoclonal anti- β -tubulin antibody as primary antibody. Rhodamine-labelled anti-mouse IgGs and FITC-labelled anti-rabbit IgGs antibodies were used as secondary antibodies.



nucleus itself^{25,26}. During this transition, DNA and microtubules would form an integrated apparatus in which DNA replication origins might have also served as microtubule-organizing centres²⁵. The origin of histones and nucleosomes has been attributed to the evolution of mitosis and intracellular cytoskeleton-based motility²⁷. In keeping with this view of evolution, our data indicate that microtubules and DNA probably use common components in their structural and functional regulation.

Many effects of histone H1 on DNA structure can be understood within a general theory of nonspecific interactions between anionic polymers and polycations²⁸. Histone H1 effects on DNA condensation can be mimicked by a variety of unrelated polycations²²; for example, in sperm nuclei DNA can be packaged either by histones, as in sea urchin, or by protamines, as in mammals²². The effects of microtubule-associated proteins on microtubule nucleation and growth can also be mimicked by unrelated and physiologically irrelevant polycations²⁹; these proteins are basic and their microtubule-stabilizing effect is probably due, at least in part, to their polycationic properties^{29,30}. The effect of the tau microtubule-associated proteins does not depend on strictly defined sites³¹. But histones have highly conserved structures and microtubule-associated proteins act as microtubule effectors *in vivo*³², so the microtubule-stabilizing and DNA-binding sequences in specific proteins must meet the requirements for correct targeting and have the appropriate regulatory sites. How is histone H1 targeted to microtubule structures? We have found no significant sequence differences between nuclear and axonemal histone H1, although histone H1 could have a covalent modification that is axoneme-specific. Perhaps histones destined for the sperm head became trapped in the adjacent tail, but the presence of histone H1 in ciliary axonemes with no association with chromatin and in *Chlamydomonas* flagella does not favour this hypothesis. The dual location of histone H1 is also a feature of some microtubule-associated proteins which can be found in nuclei³³.

Other questions concern the uniqueness of histone H1 as a microtubule-stabilizing factor in sea urchin sperm flagella and its function in the morphogenesis of flagellar structures. That this microtubule-stabilizing histone H1 is apparently unique could mean that there are other effectors which are undetectable in our assay. These assays detect proteins that are absorbed onto the surface of preformed microtubules and whose activity is not destroyed by detergent treatment, requirements that can be fulfilled by other effectors of microtubule stability like micro-

tubule-associated protein-2, tau proteins or STOP proteins⁴. Tektins can associate with stable tubulin assemblies in flagellar extracts¹⁸ and have been proposed as microtubule stabilizers, but they are inactive under our assay conditions, cannot generate microtubule stability in reconstitution experiments and are related to intermediate filament proteins^{34,35}, which do not stabilize microtubules. Maybe they need to assume an insoluble filament form³⁵ to be active, which they cannot in our assay. We find enough histone H1 in axonemes for complete microtubule stabilization, which occurs at a molar ratio of histone H1 to tubulin of ~1/10. Extraction profiles of histone H1 indicate that it might associate preferentially with the A-tubules of outer doublets, but microtubules from other axonemal structures may also be stabilized by histone H1, which would either absorb immediately onto surviving microtubules or precipitate during their disruption. The difference between A- and B-tubules with regard to stability might reflect structural differences or, as already discussed, the effect of additional factors such as tektin filaments. Although electron microscopy and immunogold labelling confirmed our immunofluorescence results, the resolution of this technique was not adequate to define exactly the location of histone H1 within different axonemal microtubules.

In our *in vitro* studies, histone H1 stabilized preformed microtubules but did not promote significant tubulin assembly, an observation which fits with the requirement for basal bodies to nucleate axonemal microtubules². We do not know yet if, in the presence of basal bodies, histone H1 can promote tubulin assembly at cold temperature. It is also unknown whether the role of histone H1 is restricted to microtubule stabilization or if it has a larger role in the morphogenesis of flagellar structures. Our finding that cilia were stained by a nuclear histone H1 affinity-purified polyclonal antibody suggests that the stabilization of complex microtubule assemblies by histone H1 is not restricted to flagella. In the scorpion *Tityus bahiensis*, histone-related proteins have also been detected in sperm axonemes using cytochemical methods³⁶. But sperm cells in other species use different proteins to package DNA; it will be interesting to see if the same is true for microtubule-stabilizing effectors.

In some cells, ciliary and flagellar microtubules are regularly disassembled during the cell cycle³⁷. Histone H1 is a major substrate for p34^{cdc2} (ref. 38), a key protein kinase in cell-cycle regulation³⁹, which raises the possibility that phosphorylation of histone H1 by this enzyme is the basis of the timely disassembly of flagellar structures during the cell cycle, an idea that is currently under investigation. □

Received 5 August; accepted 1 October 1992.

- Behrke, O. *Int. Rev. exp. Pathol.* **9**, 1-92 (1970).
- Dustin, P. *Microtubules* (Springer, New York, 1984).
- Kirschner, M. & Mitchison, T. *Cell* **45**, 329-342 (1986).
- Margolis, R. L., Rauch, C. T., Pirollet, F. & Job, D. *EMBO J.* **9**, 4095-4102 (1990).
- Farrell, K. W. *Meth. Cell Biol.* **24**, 61-78 (1982).
- Gundersen, G. G. & Bulinski, J. C. *Eur. J. Cell Biol.* **42**, 288-294 (1986).
- Sherwin, T., Schneider, A., Sasse, R., Seebeck, T. & Gull, K. *J. Cell Biol.* **104**, 439-446 (1987).
- Le Dizet, M. & Piperno, G. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5720-5724 (1987).
- Schulze, E., Asai, D. J., Bulinski, J. C. & Kirschner, M. W. *J. Cell Biol.* **105**, 2167-2177 (1987).
- Paturlle, L., Wehland, J., Margolis, R. L. & Job, D. *Biochemistry* **28**, 2698-2704 (1989).
- Stephens, R. E. *J. Cell Sci.* **101**, 837-845 (1992).
- L'Hernault, S. W. & Rosenbaum, J. L. *J. Cell Biol.* **97**, 258-269 (1983).
- Farrell, K. W. & Wilson, L. *J. molec. Biol.* **121**, 393-410 (1978).
- Pirollet, F., Rauch, C. T., Job, D. & Margolis, R. L. *Biochemistry* **28**, 835-842 (1989).
- Job, D., Pabion, M. & Margolis, R. L. *J. Cell Biol.* **101**, 1680-1689 (1985).
- Gibbons, I. R. & Frank, K. E. *J. Cell Biol.* **54**, 365-381 (1972).
- Stephens, R. E. *J. molec. Biol.* **47**, 353-363 (1970).
- Linck, R. W. *J. Cell Sci.* **20**, 405-439 (1976).
- Linck, R. W. & Stephens, R. E. *J. Cell Biol.* **104**, 1069-1075 (1987).
- Johns, E. W. *Biochem. J.* **92**, 55-59 (1964).
- Vacquier, V. D. *Meth. Cell Biol.* **27**, 15-40 (1986).
- Hnilica, L. S. *The Structure and Biological Functions of Histones* (CRC, Cleveland, 1972).
- Stollar, B. D. *Meth. Cell Biol.* **18**, 105-122 (1978).

- Cavalier-Smith, T. *Nature* **256**, 463-468 (1975).
- Cavalier-Smith, T. *Ann. N.Y. Acad. Sci.* **503**, 17-54 (1987).
- Cavalier-Smith, T. *BioEssays* **9**, 72-78 (1988).
- Cavalier-Smith, T. *Soc. gen. Microbiol.* **32**, 33-84 (1981).
- Revzin, A. *The Biology of Nonspecific DNA-Protein Interactions* (CRC, Boca Raton, 1990).
- Erickson, H. P. & Voter, W. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2813-2817 (1976).
- Harrison, A. & Hyams, J. S. *J. Cell Sci.* **96**, 347-349 (1990).
- Butner, K. A. & Kirschner, M. W. *J. Cell Biol.* **115**, 717-730 (1991).
- Drubin, D., Kobayashi, S. & Kirschner, M. W. *Ann. N.Y. Acad. Sci.* **466**, 257-268 (1986).
- Izant, J. G., Weatherbee, J. A. & McIntosh, J. *Nature* **295**, 248-250 (1982).
- Steffen, W. & Linck, R. W. *Cell Motil. Cytoskel.* **14**, 359-371 (1989).
- Stephens, R. E. *J. Cell Sci.* **92**, 403-413 (1989).
- Esponda, P. & Taffarel, M. *Cell Biol. Int. Rep.* **4**, 549-554 (1980).
- Lefebvre, P. A. & Rosenbaum, J. L. *A. Rev. Cell Biol.* **2**, 517-546 (1986).
- Labbé, J. C., Lee, M. G., Nurse, P., Picard, A. R. & Dorée, M. *Nature* **335**, 251-254 (1988).
- Nurse, P. *Nature* **344**, 503-508 (1990).
- Towbin, H., Staehelin, T. & Gordon, J. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4350-4354 (1979).
- Strickland, W. N. et al. *Eur. J. Biochem.* **104**, 567-578 (1980).
- Pirollet, F., Job, D., Margolis, R. L. & Garel, J. R. *EMBO J.* **6**, 3247-3252 (1987).

ACKNOWLEDGEMENTS. This work is dedicated to the memory of M. P. Cosson. We thank J. Cosson and P. Hutoirel for their help; Y. Saoudi and N. Scherrer for technical assistance, A. Adoutte, A. Fleury and M. H. Bré for help with immunostaining, M. Paintrand for help with electron microscopy, Y. Pierre for *Chlamydomonas* cells and N. Milloz for typing this manuscript.

Crystal structure of TFIID TATA-box binding protein

Dimitar B. Nikolov*, Shu-Hong Hu^{*†||}, Judith Lin^{*†||}, Alexander Gasch[‡], Alexander Hoffmann[§], Masami Horikoshi[§], Nam-Hai Chua[‡], Robert G. Roeder[§] & Stephen K. Burley^{*†||}

* Laboratories of Molecular Biophysics, ‡ Laboratory of Plant Molecular Biology, § Laboratory of Biochemistry and Molecular Biology and † Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA

The structure of a central component of the eukaryotic transcriptional apparatus, a TATA-box binding protein (TBP or TFIID τ) from *Arabidopsis thaliana*, has been determined by X-ray crystallography at 2.6 Å resolution. This highly symmetric α/β structure contains a new DNA-binding fold, resembling a molecular 'saddle' that sits astride the DNA. The DNA-binding surface is a curved, antiparallel β -sheet. When bound to DNA, the convex surface of the saddle would be presented for interaction with other transcription initiation factors and regulatory proteins.

EUKARYOTES have three RNA polymerases (pol I, II and III) that catalyse transcription of nuclear genes¹. Despite their structural complexity, these multisubunit enzymes require auxiliary proteins known as general initiation factors to initiate transcription from the corresponding class I, II and III promoters²⁻⁵. Until recently, it was thought that the class I, II and III initiation factors constituted non-overlapping sets of proteins. The TATA-box binding polypeptide (TBP or TFIID τ), first identified as a component of the class II initiation factor TFIID, is now known however to be required for transcription by all three nuclear RNA polymerases (reviewed in refs 6-9). It thus resembles the essential subunits common to the three RNA polymerases⁵ in being a universal transcription factor.

The role of TBP in the initiation and regulation of transcription is best understood for genes transcribed by pol II together with the general initiation factors TFIIA, -B, -D, -E, -F, -G/J, -H and -I^{3,4}. In this case, TBP is tightly associated with other polypeptides (the TAFs)¹⁰⁻¹² to form a multiprotein complex called TFIID. TFIID was first identified as a general initiation factor¹³ that binds to the TATA element and, thereby, coordinates formation of a functional preinitiation complex (PIC) by class II initiation factors and RNA polymerase II (reviewed in ref. 12). TBP cannot substitute for native TFIID in activator-dependent transcription^{14,15} (at least in higher eukaryotes), but can mediate PIC assembly¹⁶ and basal (core promoter) transcription in concert with other general class II factors. It interacts directly with the general initiation factors TFIIA and TFIIB, the C terminus of the large subunit of pol II, some cofactors that inhibit PIC formation (NC1, NC2, DR1), and an initiator-binding factor (TFII-I) that may be important for TATA-less promoters. It can also interact with certain gene-specific activators, suggesting that it may be a direct target for some of these proteins (reviewed in refs 3, 4).

The roles of TBP in transcription by pol I and pol III are less well understood. A TBP-TAF complex (SL1) has, however, been implicated in pol I transcription, and these TAFs are believed to be distinct from those in TFIID¹⁷. No pol III-specific TBP-TAF complex has yet been reported, although one is thought to exist^{8,9}.

Soon after the initial description and purification of TBP from yeast¹⁸⁻²⁰, complementary DNAs encoding TBP were cloned

from a variety of eukaryotes (Table 1). The protein has a phylogenetically conserved C-terminal portion of 180 amino acids, containing two 40% identical repeats of 66-67 amino acids flanking a highly basic segment known as the basic repeat¹⁴. This part of the protein binds to the TATA consensus sequence (TATAa/tAa/t)^{21,22} with high affinity ($k_d = 2-4 \times 10^{-9}$ M)²³, interacting with the minor groove^{24,25} and promoting DNA bending²⁶. The N-terminal part of the protein varies in length, shows little or no conservation among different organisms, and is largely unnecessary for transcription in certain yeast strains (reviewed in refs. 6,7). Although some plants have two distinct TBPs, most eukaryotes have only one.

We now report the structure of a canonical TATA-box binding protein determined by X-ray crystallography. The structure of TBP-2 from *Arabidopsis thaliana*, obtained at 2.6 Å resolution, reveals a new, highly symmetric DNA-binding fold composed of two topologically identical domains derived from the two direct repeats in the phylogenetically conserved sequence. The domains are related by an approximate 2-fold axis, and consist of a five-stranded, antiparallel β -sheet and two α -helices. The intramolecular symmetry generates a saddle-shaped structure dominated by a curved antiparallel β -sheet, which forms the saddle's concave face and interacts with the DNA. The convex face of the saddle would interact with other proteins during transcription.

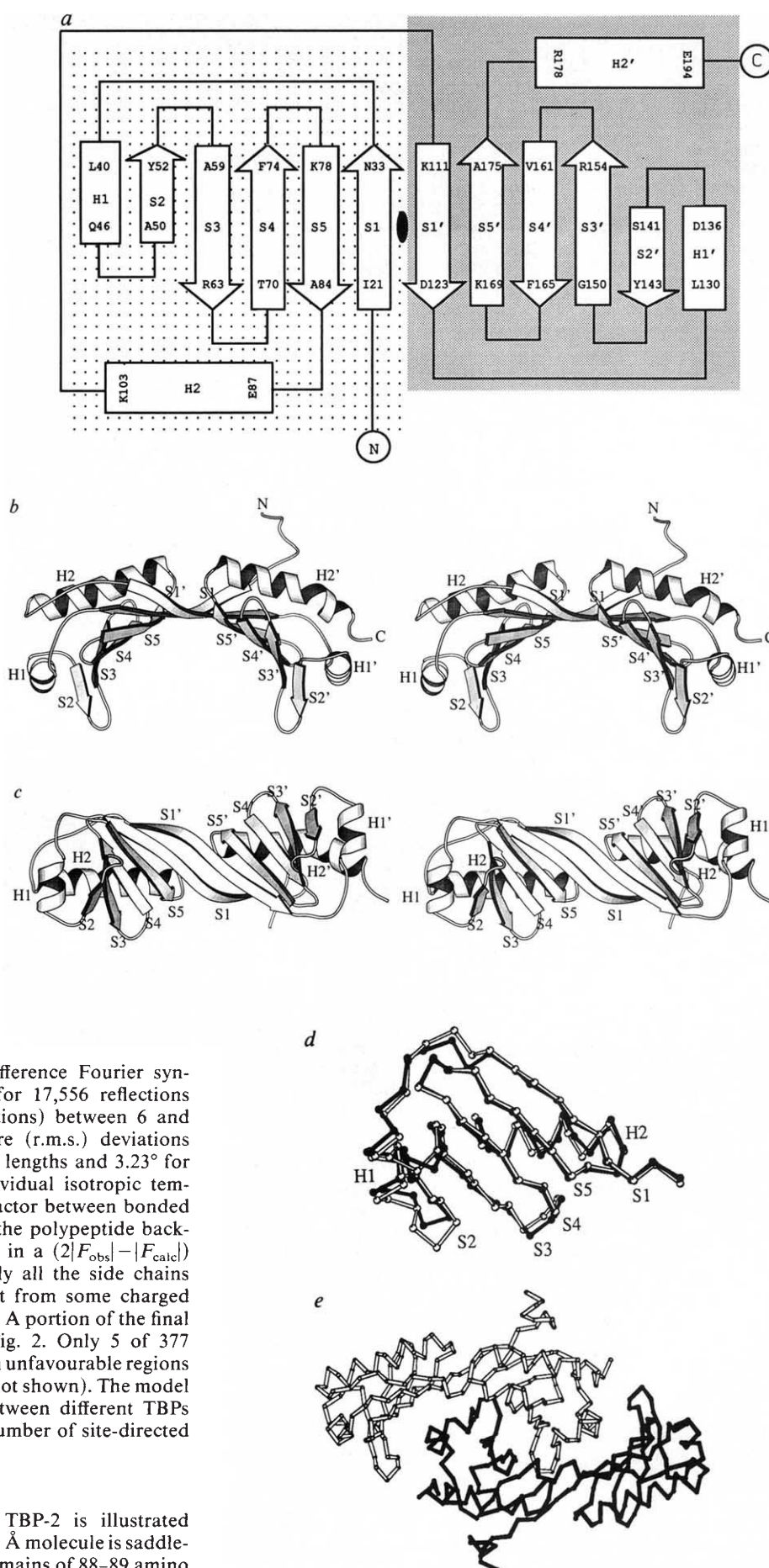
Crystallization and structure determination

TBP-2 from *Arabidopsis thaliana*²⁷ was expressed and purified, and diffraction-quality tetragonal crystals (space group I4₂2₂; $a = 105.1$ Å, $c = 215.6$ Å) with two molecules per asymmetric unit were obtained as described in Table 2. The structure was solved by multiple isomorphous replacement (MIR) using five mercurial derivatives (Table 2). The 3 Å resolution MIR map revealed clear electron density for both α -helices and β -strands. After 'solvent flattening'²⁸ a partial, discontinuous poly-alanine model was fitted to the modified electron density, and the non-crystallographic symmetry operator and molecular boundaries were determined. The electron density map was further improved by phase combination of the refined partial structure model at 3.0 Å resolution²⁹ with the MIR phase data³⁰. Two rounds of model building³¹, refinement and phase combination gave an unambiguous trace for about 93% of the asymmetric unit. The resolution of the native data was extended to 2.58 Å using simulated annealing and further refinement. The current model consists of residues 7 to 198 and 13 to 199 in the first and second monomers. No water molecules or sulphate anions have been included in the model, despite appropriately

|| To whom correspondence should be addressed.

† Present addresses: St Vincent's Institute of Medical Research, Fitzroy, Victoria 3065, Australia (S.-H.H.); Department of Medicine, University of Southern California Medical Center, Los Angeles, California 90033-1084, USA (J.L.).

FIG. 1 Structure of TBP-2 from *Arabidopsis thaliana* determined by X-ray crystallography. Residues were assigned to secondary structural elements according to criteria defined by Kabsch and Sander⁵⁰. The N and C termini of the protein are labelled, as are the residues at the end of each numbered α -helix (H) and β -strand (S). The symbol ' refers to the second structural domain. *a*, Schematic drawing of the secondary structure. α -Helices are shown as rectangles and β -strands as shown as arrows. The location of the axis of approximate 2-fold symmetry is indicated by the symbol (●). The boundaries of the two structural domains are depicted with shading. *b*, Stereodrawing of a Richardson-type⁵¹ representation of the three-dimensional structure viewed perpendicular to the intramolecular 2-fold symmetry axis. The symmetry breaks down slightly in helix H2', where Pro190 induces a bend. *c*, Stereodrawing viewed along the intramolecular symmetry axis, showing the β -sheet structure on the underside of the molecular saddle. *d*, Superposition of the α -carbon backbone of the first and second domains comprising TBP-2. α -Carbon atomic positions of residues 24–46 and 49–98 were used to determine the least-squares best superposition of the two domains, depicted with open and solid lines (r.m.s. deviation of 1.0 Å for the α -carbon atomic pairs). Significant differences occur at the loop connecting helix H1 with strand S2, where there is a single amino-acid insertion in the second domain (His 137). The two domains are about 30% and 50% identical at the amino-acid and nucleotide levels, respectively. *e*, Schematic drawing of the crystallographic asymmetric unit showing the α -carbon backbone trace of the two tightly associated molecules.



positioned features in ($|F_{\text{obs}}| - |F_{\text{calc}}|$) difference Fourier syntheses. The current *R*-factor is 23.0% for 17,556 reflections ($|F| > 1\sigma(|F|)$; 99% of possible observations) between 6 and 2.58 Å resolution, with root-mean-square (r.m.s.) deviations from ideal geometry of 0.015 Å for bond lengths and 3.23° for bond angles, and tightly restrained individual isotropic temperature factors (r.m.s. deviations in *B*-factor between bonded atoms = 2.3 Å²). The electron density of the polypeptide backbone is everywhere continuous at 1.25 σ in a ($2|F_{\text{obs}}| - |F_{\text{calc}}|$) difference Fourier synthesis, and virtually all the side chains have well defined electron density, apart from some charged residues located on the molecular surface. A portion of the final electron density map is illustrated in Fig. 2. Only 5 of 377 backbone torsion angle combinations lie in unfavourable regions of the Ramachandran (ϕ , ψ) plot³² (data not shown). The model is also supported by the comparison between different TBPs described below and the locations of a number of site-directed mutations.

Structure description

Topological overview. The structure of TBP-2 is illustrated schematically in Fig. 1*a–c*. The 32 × 45 × 60 Å molecule is saddle-shaped with two very similar structural domains of 88–89 amino

TABLE 1 TBP sequence alignment

Mutations			10																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
-----------	--	--	----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Ten TBP sequences, aligned by sequence homology^{14,15,27,38,52-62}. The first sequence block includes the entire nonconserved N-terminal region of *Arabidopsis*, maize, and potato TBPs, and the corresponding 18 residues from the TBPs of the remaining organisms. The second and third blocks contain each of the two structural repeats comprising the conserved C-terminal region. Amino acids occupying identical positions in each repeat are aligned vertically. Only differences from the *Arabidopsis* TBP-2 sequence are shown. A gap is present in the first domain to accommodate the insertion of a single amino acid in the second. Secondary structural elements are labelled and indicated with S for strand and H for helix, and the long arrow denotes the extent of the direct repeats. The locations of point mutants and their phenotype are indicated as: Mutations affecting: * DNA binding, † All function, ‡ DNA recognition, § TFIIA interaction, || SPT3 interaction, ¶ Pol. specificity, # No effect.

acids related by approximate intramolecular 2-fold symmetry. Each α/β domain is comprised of two α -helices and a five-stranded, antiparallel β -sheet connected in the order S1-H1-S2-S3-S4-S5-H2. The spatial order of the five β -strands is S1-S5-S4-S3-S2, with the short H1 helix on the edge of the sheet adjacent to strand S2 and the long, amphipathic H2 helix abutting the convex side of the sheet and defining the domain's hydrophobic core. The two domains are connected by a short polypeptide derived from the basic repeat (subsequently referred to as the basic linker peptide).

The loops connecting strands S2 and S3 and S2' and S3', respectively, can be visualized as the 'stirrups' of the molecular saddle. The underside of the saddle forms a half-cylindrical recess 32×25 Å, lined by the central 8 strands of the 10-stranded β -sheet, which is generated by 2-fold symmetry between the 5-stranded β -sheets of each structural domain. Solvent-accessible residues on the underside include the hydrophobic and small uncharged polar side chains projecting from each of the

eight β -strands and the positively-charged residues located on the six loops connecting the strands (Fig. 3a). The convex surface of the saddle is composed of the four α -helices, the basic linker peptide, parts of strands S1 and S1', and the 18 nonconserved N-terminal residues. To the best of our knowledge, the structure of TBP-2 represents a new protein fold.

Intramolecular 2-fold symmetry. The two structural domains of TBP-2 are topologically identical (Fig. 1d). Although this finding was not entirely unexpected given the similarity of the direct repeats, the three-dimensional structural overlap goes well beyond the repeats, encompassing almost all of the conserved C-terminal domains. The structural repeats in TBP thus include much of the basic segment between the two direct repeats and the extreme C terminus of TBP (as hypothesized by Stucka and Feldmann³³).

Non-crystallographic symmetry. The two molecules forming the crystallographic asymmetric unit are related by a non-crystallographic 2-fold axis of rotation (Fig. 1e). The r.m.s. deviation

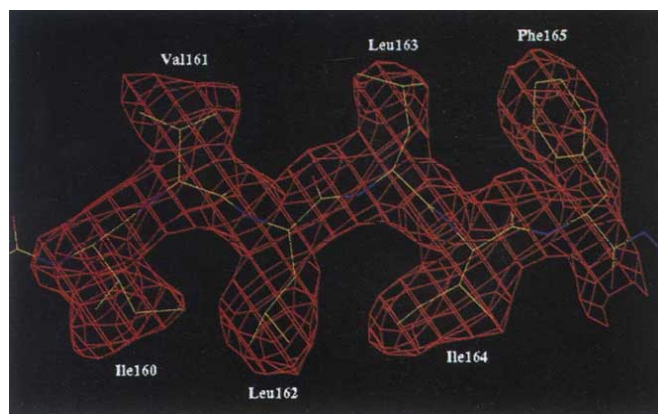


FIG. 2 Electron density map of TBP-2. View of the region of the structure shown to determine the specificity for TATA-box binding. The $(|F_{\text{obs}}| - |F_{\text{calc}}|)$ Fourier difference synthesis was calculated at 2.58 Å resolution using phases calculated by omitting residues 160 to 165 from the refined structural model. The contour level is 2.5σ and the refined atomic model is shown as a stick figure.

between α -carbon atomic positions of the two is 0.48 Å for residues 13 to 198. The intermolecular packing appears to result from tight self-association of TBP. Hydrodynamic studies of various TBPs show that they dimerize readily in solution, unless complexed with DNA containing the TATA-box (see legend to Table 2).

Sequence comparison of TBPs

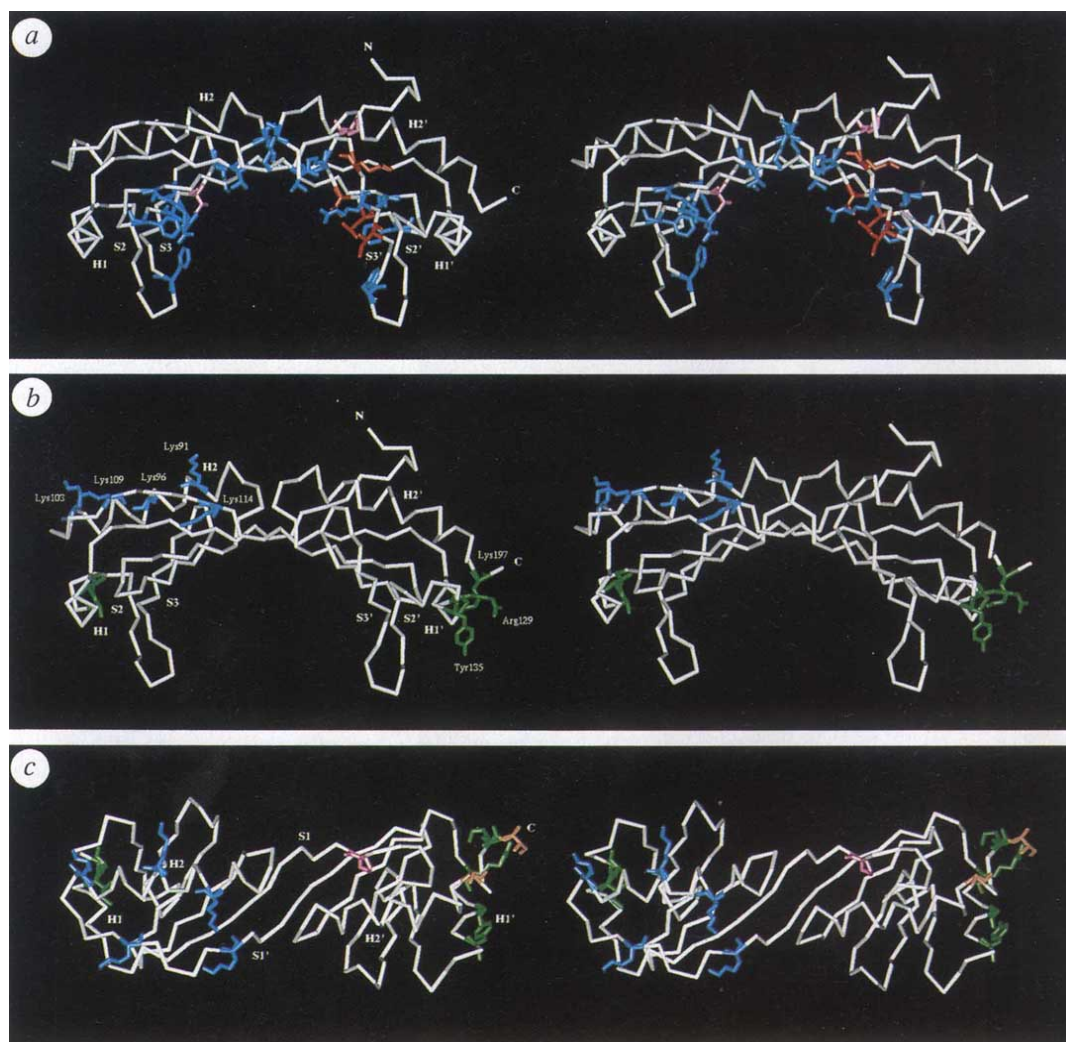
Seventy per cent of the 180 amino acids in the C-terminus of TBP are conserved in all species so far examined (Table 1). All 54 sites at which differences occur map either to the surface of

the molecule, where mutations would be tolerated, or represent conservative changes of buried residues unlikely to destabilize the protein's hydrophobic core. All known TBPs must thus share the same three-dimensional structure in this region³⁴. Subsequent references to the results of mutagenesis studies use *Arabidopsis* TBP-2 sequence numbers throughout (see Table 1 for conversion to other sequence numbering schemes).

Species specificity in TBP is largely determined by the conserved 180 amino acids^{35,36}. In yeast-human hybrid molecules, residues 130–155 (helix H1' and strands S2' and S3') and 173–188 (helix H2') must come from yeast to support normal growth in *Saccharomyces cerevisiae*. The only differences between the yeast and human sequences lie in helix H1', the loop connecting helix H1' and strand S2', and helix H2'. These regions constitute part of the convex protein-binding surface in TBP (see below).

The 18 N-terminal amino acids of TBP-2 have not yet been visualized in their entirety, as the two copies in the asymmetric unit adopt distinct conformations and show evidence of disorder. This portion of the TBP sequence is conserved only among plants (Table 1). Normal TBP function in *S. cerevisiae* requires

FIG. 3 Schematic stereo-drawings of selected TBP point mutants. The α -carbon backbone is shown as a solid white line and highlighted residues are drawn as colour-coded atomic stick figures. *a*, Residues implicated in DNA binding and RNA polymerase specificity, displayed using the view in Fig. 1*b*. Red, alteration in TATA-box specificity; blue, reduced affinity for DNA; pink, alteration in RNA polymerase specificity. *b*, Viewed as in Fig. 1*b*. Residues implicated in TBP-transcription factor interactions: blue, TBP-TFIIA interaction; green, TBP-SPT3 interaction. *c*, View of the upper surface of the molecular saddle. Ser 193 and Ile 198, which may distinguish TBP isoform 1 from TBP isoform 2 in *Arabidopsis* are shown in brown.



several acidic amino acids in this region, but is insensitive to their precise locations^{37,38}. Both *Schizosaccharomyces pombe* and plants also possess a relative abundance of acidic residues in this region.

DNA binding and promoter recognition

Biochemical and genetic data confirm that the TBP saddle sits astride the DNA (Figs 3a and 4). The model in Fig. 4 protects ~10 base pairs, in good agreement with the 15 base pair DNaseI footprint of *S. cerevisiae* TBP²⁰.

The locations of point mutations reducing DNA-binding affinity, altering TATA-binding specificity, and changing polymerase specificity are shown in Fig. 3a. All mutations affecting DNA binding map to the underside or edges of the molecular saddle. The different mutations may eliminate positive charges from the loops connecting segments of the β -sheet, disrupt the structure of the stirrup-like loops between strands S2 and S3 or S2' and S3', or introduce positively charged residues in place of the nonpolar or small, uncharged polar side chains located on the β -strands of the saddle^{21,22,39}.

Discussion of the TBP-TATA-box interaction must await high-resolution structures of appropriate protein-DNA complexes. Mutations affecting TATA-box binding and polymerase

specificity are nevertheless of considerable interest. Two amino-acid changes, Ile 152 $\rightarrow$ Phe and Leu 163 $\rightarrow$ Val (Fig. 3a, red side chains), are necessary and primarily responsible for changing DNA-binding specificity from TATA to TGTA in yeast³⁹. These substitutions map to strands S3' and S4', respectively, where they would alter the precise shape of the DNA-binding surface. The mutations Lys 156 $\rightarrow$ Gln/Asn and Val 161 $\rightarrow$ Thr, contribute to the effect and map to the loop connecting strands S3' and S4' and to strand S4', respectively. Two further mutations⁷ affect the polymerase specificity of TBP (Fig. 3a, pink side chains). Thr 70 $\rightarrow$ Lys, which reduces DNA-binding affinity and abolishes pol II activity, maps to strand S4 and would change the structure and electrostatic appearance of TBP's DNA-binding surface. Pro 23 $\rightarrow$ Ser, which abolishes both pol II and pol III activities, maps to the beginning of strand S1 where a mutation could alter the convex protein-binding surface of TBP (see below and Fig. 3b, c).

The half-cylindrical recess defined by the concave surface of TBP-2 will accommodate B-form DNA without steric overlap in computer-generated models (Fig. 4a). Unlike any other structurally characterized DNA-binding protein⁴⁰⁻⁴² TBP recognizes its binding site through specific contacts with the minor groove^{24,25}. The DNA-binding face of the protein is a curved, eight-stranded β -sheet, which provides a large surface area for DNA contact. This area may also mediate nonspecific DNA-protein contacts required to position TBP on TATA-deficient promoters for functions analogous to those required during transcription from TATA-containing promoters⁴³. Unfortunately, our data do not fix the relative orientations of TBP and the DNA to which it binds, or the relative location of the transcription start site. The DNA-binding surface may flex about a hinge between TBP's two structural domains.

TBP-transcription factor interactions

The crystal structure of TBP-2 also provides a basis for analysing contacts between it and the other proteins involved in transcription, such as general transcription initiation factors, promoter-specific factors and viral transactivators.

Five lysine residues in the basic repeat of TBP are involved in the interaction with TFIIA^{44,45} a general transcription initiation factor that binds to and stabilizes the TBP-TATA complex during assembly of the PIC¹⁶. (Fig. 3b, c; Table 1). They map to a highly basic area of the convex surface of TBP, containing the polar face of helix H2 and the basic linker. Non-conservative point mutations at any of these sites reduce the strength of the TBP-TFIIA interaction, which is presumably electrostatic.

Four sites on the surface of TBP, where the helix H1' and the C terminus of the polypeptide chain are in close proximity, are involved in the interaction between TBP and a *S. cerevisiae* protein SPT3, a specificity factor required for transcription initiation at certain promoters (Fig. 3b, c). The wild-type amino acids include Arg 129, Gly 132, Tyr 135 (Phe in yeast), and Lys 197 (Table 1). Specific polar and nonpolar contacts between the molecular surfaces of TBP and SPT3 are therefore probably responsible for stabilizing this complex. A double mutation at positions 43 and 44, located on the solvent-inaccessible face of helix H1 (Fig. 3b, c), implicates another part of TBP's surface in interactions with SPT-3.

The *in vitro* association of the adenovirus transactivator E1A with TBP depends on amino acid(s) located between Lys 85 and Tyr 135^{47,48}. This region of TBP encompasses helix H2, the basic linker connecting the domains of TBP, strand S1' and the majority of helix H1', all of which contribute to TBP's protein-binding surface. Consequently, the E1A binding site may include residues that also mediate the TBP-TFIIA interaction.

Many other TBP-protein interactions occur during transcription initiation, and the protein-binding surface of TBP is by no means fully characterized. The binding sites for TFIIB, the C terminus of the large subunit of pol II, the polymerase-specific TAFs, negative cofactors, and transcriptional activator proteins

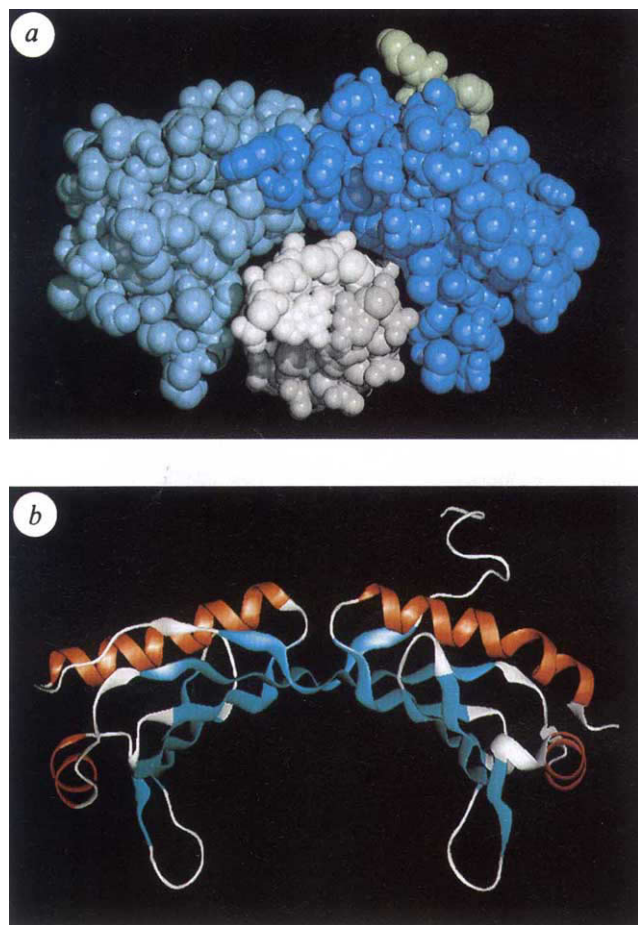


FIG. 4 a, Computer-generated model of TBP-2 interacting with B-form DNA without steric repulsion. The protein atoms are colour-coded by their location within the structure (light green, non-conserved segment of the polypeptide chain; light blue, first domain; dark blue, second domain), and the atoms of each DNA strand are colored white and grey, respectively. Each atom is depicted as a sphere of radius equal to the atomic van der Waals radius. It should be emphasized that this figure does not represent the TBP-DNA interaction with any precision, and is included only to illustrate the general picture of DNA-binding derived from this structural study. b, Ribbon drawing of TBP-2 viewed as in a. α -Helices and β -strands are shown in red and blue, respectively, with the remainder in white.

TABLE 2 Summary of crystallographic analysis

	Native	THM	MCL	EMP	DMM	MMC
Resolution (Å)	2.58	2.70	2.60	2.70	3.00	3.00
Reflections (observed/unique)	138,216/19,227	49,979/16,157	105,493/18,460	40,469/13,727	36,605/10,834	44,696/10,215
Data coverage (%)	99	94	97	80	86	82
R_{sym} (%)	6.8	5.9	7.8	8.7	8.1	7.2
Mean fractional isomorphous difference		0.29	0.33	0.16	0.24	0.23
MIR analysis (10–3 Å)						
Number of sites		5	9	5	6	5
Phasing power		1.73	1.82	1.16	1.45	1.44
R_c		0.59	0.60	0.64	0.58	0.61
Mean overall figure of merit	0.64					
Refinement						
Resolution (Å)	6–2.58					
R -factor (%)	23.0					
Reflections ($ F > \tau F $)	17,556					
Total number of atoms	2,988					
r.m.s. bond length (Å)	0.015					
r.m.s. bond angle (degrees)	3.23					

TBP-2 derived from an AT-2 cDNA from *Arabidopsis thaliana*²⁷ fused with an additional 20 N-terminal amino acids (MGSSHHHHSSGLVPRGSH; single-letter code) was overexpressed in *Escherichia coli* (BL21(DE3)pLysS) using the T7 RNA polymerase system^{63,64}. Cells grown at 30 °C to an absorbance of 0.7 at 595 nm and induced with 0.5 mM isopropyl- β -D-thiogalactopyranoside for 3 h were collected by low-speed centrifugation. Lysis was done by three cycles of freeze-thaw and DNA digestion was done with DNaseI. The soluble fraction was applied directly to a Ni²⁺-ion affinity column and washed with a buffer containing increasing amounts of imidazole. TBP with an estimated homogeneity of 95% was then eluted from the resin with a buffer containing 100 mM EDTA. After overnight dialysis, removal of the histidine-containing N-terminal sequence was done by either thrombin or trypsin proteolytic cleavage of the peptide bond between R and G in the above sequence to yield TBP-2 plus three N-terminal amino acids, namely GS and H. A final cation-exchange chromatography step to remove the protease and any remaining contaminants yielded material of homogeneity estimated to be 98% or greater. Mass spectrometry⁶⁵ documented that the TBP used for crystallization was neither modified nor further proteolysed during expression and purification. The measured molecular mass was $22,650 \pm 5$, compared with 22,653 calculated from the predicted amino-acid sequence. Further control experiments were done to ensure that the protein used for the crystallographic study was biochemically active and indistinguishable from recombinant yeast TBP. Gel retardation analyses, done as described in ref. 20, confirmed that the purified protein had full DNA-binding activity (data not shown). Photon correlation spectroscopy (PCS) documented that TBP is a dimer in solution in the absence of DNA and undergoes a dimer-to-monomer transition on binding to oligonucleotides containing the TATA-consensus sequence. The PCS experiments were done with a Biotage dp801 molecular size detector under a wide range of experimental conditions. Purified, recombinant *Arabidopsis thaliana* and *S. cerevisiae* TBPs were monodisperse and dimeric in the absence of DNA. Addition of duplex oligonucleotide containing the TATA-consensus sequence yielded a TBP-DNA complex consisting of 1 molecule of TBP and 1 molecule of duplex oligonucleotide. Presumably, the crystallographic asymmetric unit corresponds to the TBP dimer detected by PCS. Crystals were obtained from 2.0 M ammonium sulphate in a buffer consisting of 20 mM HEPES pH8.5 and 5 mM MgCl₂ using vapour diffusion combined with microscopic seeding. X-ray data collection at conventional sample temperatures (4–20 °C) was complicated by a high degree of variation in the unit cell dimensions between crystals grown under identical conditions, and during the course of measurements from a single crystal maintained at constant temperature. Diffraction experiments were carried out at -150 °C using a Rigaku RAXIS-II imaging plate area detector and a refrigerated nitrogen gas delivery system. Crystals were transiently 'cryoprotected' in a 30% glycerol, 15% PEG4K mixture before placement in a 1 mm diameter wire loop and freezing. Significant unit cell edge length fluctuations were eliminated by optimization and standardization of the freezing protocol, and the unit cell dimensions did not vary by more than ± 0.3 Å. Heavy-atom derivatives were prepared by soaking the crystals with mercurial reagents at concentrations of 0.1–1.0 mM for 15–60 h (DMM, dimercy malonate; EMP, ethyl mercury phosphate; MMC, methyl mercury chloride; MCL, mercuric chloride; THM, thimerosal). All derivative data were measured from single crystals, whereas the native data were measured from two crystals. There was no evidence of non-isomorphism between the two native crystals, and no significant non-isomorphism between native and derivative crystals. Oscillation photographs were integrated and reduced to structure factor amplitudes using software provided by the manufacturer. Additional crystallographic software included the PROTSYS package (G. A. Petsko, personal communication), HEAVY⁶⁶, and MOLSCRIPT⁶⁷. To our knowledge this is the first MIR structure determination undertaken entirely using crystals frozen at near liquid nitrogen temperature. Atomic coordinates and structure factor amplitudes have been submitted to the Brookhaven Protein Data Bank⁶⁸. $R_{\text{sym}} = \sum |I - \langle I \rangle| / \sum I$, where I = observed intensity, $\langle I \rangle$ = average intensity obtained from multiple observations of symmetry related reflections. Mean fractional isomorphous difference = $\sum ||F_{\text{PH}}| - |F_{\text{P}}|| / \sum |F_{\text{P}}|$, $|F_{\text{PH}}|$ = protein structure factor amplitude, $|F_{\text{P}}|$ = heavy-atom derivative structure factor amplitude. $R_c = \sum ||F_{\text{H(obs)}}| - |F_{\text{H(calc)}}|| / \sum |F_{\text{H(obs)}}|$, $|F_{\text{H(obs)}}|$ = observed heavy atom structure factor amplitude and $|F_{\text{H(calc)}}|$ = calculated heavy-atom structure factor amplitude. Phasing power = root mean square ($|F_{\text{H}}|/E$), $|F_{\text{H}}|$ = heavy-atom structure factor amplitude and E = residual lack of closure. r.m.s. bond lengths and r.m.s. bond angles are the respective r.m.s. deviations from ideal values.

remain undetermined.

TBP diversity in plants

Unlike mammals, *Drosophila* and yeast, *Arabidopsis thaliana* and maize each have two different TBPs (reviewed in ref. 49). Comparison of the plant sequences reveals only two positions at which both pairs of proteins differ (Table 1). In TBP-1, residue 193 is arginine in *Arabidopsis* and alanine in maize, whereas in both TBP-2s it is serine. Residue 198 is valine in both TBP-1s and isoleucine in both TBP-2s. It is unknown whether TBP-1 and TBP-2 have identical roles in *Arabidopsis* and in maize, but because residues 193 and 198 occur on the surface of helix H2' (Fig. 3c), functional differences between the two TBP isoforms may result from isoform-specific interactions with other proteins.

Conclusion

TBP is a saddle-shaped protein with a new DNA-binding fold. The concave DNA-binding surface of the saddle is a curved antiparallel β -sheet, which may mediate both specific and non-

specific contacts with DNA. On binding to DNA, the convex surface of the saddle would be presented for interaction with TAFs, transcription initiation factors and other regulatory proteins. Although the structure of the protein displays approximate intramolecular 2-fold symmetry, neither the DNA- nor the protein-interaction surfaces of TBP are symmetric, and the two domains present chemically distinct patterns of amino-acid side chains. TBP can thus bind DNA and other proteins in a directional manner.

The structure presented here provides a starting point for further crystallographic, biochemical and genetic studies of TBP. In particular, it should now be possible systematically to determine first, which regions of TBP's protein-binding surface mediate very tight (and possibly irreversible) interactions with polymerase-specific TAFs, and second, how the TBP-TAF complex (or in yeast TBP itself) interacts reversibly with other proteins involved in transcription. Given the somewhat limited protein-binding surface of TBP, many of the interaction sites may be overlapping and competition for them

may influence transcription.

Finally, the approximate intramolecular symmetry and amino-acid sequence similarity of TBP's two structural domains sug-

gests that the primordial ancestor of TBP functioned as a dimeric DNA-binding protein before gene duplication and fusion produced the structure described here. □

Received 10 September; accepted 19 October 1992.

1. Sentenac, A. *CRC Curr. Rev. Biochem.* **18**, 31–90 (1985).
2. Reeder, R. H. *Trends Genet.* **6**, 390–395 (1990).
3. Roeder, R. G. *Trends biochem. Sci.* **16**, 402–408 (1991).
4. Zawel, L. & Reinberg, D. *Curr. Opin. Cell Biol.* **4**, 488–495 (1992).
5. Gabrielsen, O. S. & Sentenac, A. *Trends biochem. Sci.* **16**, 412–416 (1991).
6. Cormack, B. P. & Struhl, K. *Cell* **69**, 685–696 (1992).
7. Schultz, M. C., Reeder, R. H. & Hahn, S. *Cell* **69**, 697–702 (1992).
8. Sharp, P. A. *Cell* **68**, 819–821 (1992).
9. White, R. J. & Jackson, S. P. *Trends Genet.* **8**, 284–288 (1992).
10. Dynlacht, B. D., Hoey, T. & Tjian, R. *Cell* **66**, 563–576 (1991).
11. Pugh, B. F. & Tjian, R. *Genes Dev.* **5**, 1935–1945 (1991).
12. Takada, R. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
13. Matsui, T., Segall, J., Weil, P. A. & Roeder, R. G. *J. biol. Chem.* **255**, 11992–11996 (1980).
14. Hoffmann, A. *et al. Nature* **346**, 387–390 (1990).
15. Peterson, M. G., Tanese, N., Pugh, B. F. & Tjian, R. *Science* **248**, 1625–1630 (1990).
16. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. *Cell* **56**, 549–561 (1989).
17. Comai, L., Tanese, N. & Tjian, R. *Cell* **68**, 965–976 (1992).
18. Cavallini, B. *et al. Nature* **334**, 77–80 (1988).
19. Buratowski, S., Hahn, S., Sharp, P. A. & Guarente, L. *Nature* **334**, 37–42 (1988).
20. Horikoshi, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 4843–4847 (1989).
21. Yamamoto, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 2844–2848 (1992).
22. Reddy, P. & Hahn, S. *Cell* **65**, 349–357 (1991).
23. Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5718–5722 (1989).
24. Lee, D. K., Horikoshi, M. & Roeder, R. G. *Cell* **67**, 1241–1250 (1991).
25. Starr, D. B. & Hawley, D. K. *Cell* **67**, 1231–1240 (1991).
26. Horikoshi, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 1060–1064 (1992).
27. Gasch, A., Hoffmann, A., Horikoshi, M., Roeder, R. G. & Chua, N.-H. *Nature* **346**, 390–394 (1990).
28. Wang, B.-C. *Meth. Enzym.* **115**, 90–112 (1985).
29. Bringer, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
30. Kabsch, W., Mannherz, H. G., Suck, D., Pai, E. F. & Holmes, K. C. *Nature* **347**, 37–44 (1990).
31. Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr.* **A47**, 110–119 (1991).
32. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. *J. molec. Biol.* **7**, 95–99 (1963).
33. Stucka, R. & Feldmann, H. *FEBS Lett.* **261**, 223–225 (1990).
34. Sander, C. & Schneider, R. *Proteins Struct. Funct. Genet.* **9**, 56–68 (1991).
35. Cormack, B. P., Strubin, M., Ponticelli, A. S. & Struhl, K. *Cell* **65**, 341–348 (1991).
36. Gill, G. & Tjian, R. *Cell* **65**, 333–340 (1991).
37. Zhou, Q., Schmidt, M. C. & Berk, A. J. *EMBO J.* **10**, 1843–1852 (1991).
38. Fikes, J. D., Becker, D. M., Winston, F. & Guarente, L. *Nature* **346**, 291–294 (1990).
39. Strubin, M. & Struhl, K. *Cell* **68**, 721–730 (1992).

40. Harrison, S. C. *Nature* **353**, 715–719 (1991).
41. Pabo, C. O. & Sauer, R. T. *A. Rev. Biochem.* **61**, 1053–1095 (1992).
42. Phillips, S. E. V. *Curr. Opin. struct. Biol.* **1**, 89–98 (1991).
43. Wiley, S. R., Kraus, R. J. & Mertz, J. E. *Proc. natn. Acad. Sci. U.S.A.* **89**, 58144–5818 (1992).
44. Lee, D. K., DeJong, J., Hashimoto, S., Horikoshi, M. & Roeder, R. G. *Molec. cell. Biol.* **12**, 5189–5196 (1992).
45. Ranish, J. A., Lane, W. S. & Hahn, S. *Science* **255**, 1127–1129 (1992).
46. Eisenmann, D. M., Arndt, K. M., Ricupero, S. L., Rooney, J. W. & Winston, F. *Genes Dev.* **6**, 1319–1331 (1992).
47. Horikoshi, N. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 5124–5128 (1991).
48. Lee, W. S., Kao, C. C., Bryant, G. O., Liu, X. & Berk, A. J. *Cell* **67**, 365–376 (1991).
49. Katagiri, F. & Chua, N.-H. *Trends Genet.* **8**, 22–27 (1992).
50. Kabsch, W. & Sander, C. *Biopolymers* **22**, 2577–2637 (1983).
51. Richardson, J. S. *Meth. Enzym.* **115**, 359–390 (1985).
52. Haass, M. M. & Feix, G. *FEBS Lett.* **301**, 294–298 (1992).
53. Holdsworth, M. J., Grierson, C., Schuch, W. & Bevan, M. *Pl. molec. Biol.* **19**, 455–464 (1992).
54. Kawata, T., Minami, M., Tamura, T.-A., Sumita, K. & Iwabuchi, M. *Pl. molec. Biol.* **19**, 867–872 (1992).
55. Hoffmann, A. *et al. Genes Dev.* **4**, 1141–1148 (1990).
56. Horikoshi, M. *et al. Nature* **341**, 299–303 (1989).
57. Cavallini, B. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9803–9807 (1989).
58. Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. *Cell* **58**, 1173–1181 (1989).
59. Schmidt, M. C., Kao, C. C., Pei, R. & Berk, A. J. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7785–7789 (1989).
60. Hoey, T., Dynlacht, B. D., Peterson, M. G., Pugh, B. F. & Tjian, R. *Cell* **61**, 1179–1186 (1990).
61. Muhich, M. L., Iida, C. T., Horikoshi, M., Roeder, R. G. & Parker, C. S. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9148–9152 (1990).
62. Kao, C. C. *et al. Science* **248**, 1646–1650 (1990).
63. Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. *Meth. Enzym.* **185**, 60–89 (1990).
64. Hoffmann, A. & Roeder, R. G. *Nucleic Acids Res.* **19**, 6337 (1991).
65. Chait, B. T. & Kent, S. B. H. *Science* **257**, 1885–1894 (1992).
66. Terwilliger, T. & Eisenberg, D. *Acta Crystallogr.* **A39**, 813–817 (1983).
67. Kraulis, J. P. *J. appl. Crystallogr.* **24**, 946–950 (1991).
68. Bernstein, F. C. *et al. J. molec. Biol.* **112**, 535–542 (1977).

ACKNOWLEDGEMENTS. D.B.N. thanks A. Ferre-D'Amaré for help and advice throughout this work and we thank S. Jacques for help with protein production, B. Chait and S. Cohen for doing the mass spectrometry at the National Resource for Mass Spectrometric Analysis of Biological Macromolecules and R. Brown, K. L. Clark, X. P. Kong, J. Kuriyan, P. Model and G. A. Petsko for their many useful suggestions. A.G. was supported by postdoctoral fellowships from the Deutsche Forschungsgemeinschaft and the Norman and Rosetta Winston Foundation. A.H. is a Boehringer-Ingelheim Graduate Fellow. M.H. was an Alexandrine and Alexander L. Sinsheimer Scholar. This work was supported by the Howard Hughes Medical Institute (S.K.B.) and by the NIH (M.H. and R.G.R.), Monsanto (N.-H.C.) and the Pew Trust.

LETTERS TO NATURE

Possible cataclysmic variable in the core of the globular cluster 47 Tucanae

Francesco Paresce^{*†}, Guido De Marchi^{*†}
& Francesco R. Ferraro[‡]

^{*} Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA

[†] Osservatorio Astronomico di Bologna, Via Zamboni 33, 40126, Bologna, Italy

SEVERAL globular clusters have low-luminosity X-ray sources at their cores^{1–3}, of which X0021.8–7221 in 47 Tucanae (NGC104) is one of the brightest. The nature of these sources is a mystery, and attempts to detect optical counterparts have been frustrated by the overcrowding of stars in cluster cores. The resolution of ground-based observations is insufficient to pick out the faint blue objects that are likely to be the optical counterparts to the X-ray sources, but the Faint Object Camera (FOC) on the Hubble Space Telescope has proved effective in identifying blue objects in dense clusters such as 47 Tuc⁴. Thus encouraged to use the FOC to search for the optical counterpart to X0021.8–7221, we have discovered a faint, variable and very blue object located in the error circle⁵ for X0021.8–7221 from the Einstein satellite High-Resolution Imager. The object has the spectral characteristics of

a cataclysmic variable, and the high inferred X-ray to optical brightness ratio suggests that it is a magnetic cataclysmic variable, one of the candidates that has been suggested as the X-ray source in globular clusters.

Several working hypotheses have been put forward for the nature and origin of the low-luminosity X-ray sources in globular clusters. They were at first thought to be normal cataclysmic variables (CVs) created in abundance in the dense cores by tidal capture between a red star and a white dwarf^{7,8}. Reconsideration of this process, as there is no evidence that they are present in such large numbers (the dwarf nova V101 in M5 is probably the only confirmed CV in a globular cluster⁹), has shown¹⁰ that its usual result might be unstable mass transfer, sharply reducing the population of CVs. Other possibilities, then, include magnetic CVs⁶, low-mass X-ray binaries viewed at high inclination¹¹, soft X-ray transients in quiescence¹² and close binaries in which there is a rapidly rotating secondary with hotspots^{10,13}.

To distinguish between these possibilities and to shed light on the nature of close binaries that may effect the dynamical evolution of globular cluster cores during and after collapse¹⁴, ultraviolet or optical identification of a counterpart is essential. We have carefully searched the field contained in the High-Resolution Imager (HRI) error circle overlapping with the available FOC images to look for objects with the markedly blue spectrum expected from a CV accretion disk. The investigation revealed within the small HRI error circle an interesting object, listed as DPF-2213 in the table provided by De Marchi *et al.*¹⁵, which we shall call V1 for simplicity (Fig. 1). The F342W and F195W filters used here are described in ref. 16. The combination of these two filters has a peak response at 3,400 Å and a full width at half maximum of 706 Å, approximating the standard Johnson U bandpass. The instrumental arrangement in this

[†] F.P. is on assignment from the Astrophysics Division, Space Science Department, ESA, and on leave from the Osservatorio Astronomico di Torino, Italy. G.D.M. is on leave from the Dipartimento di Astronomia, Scuola di Dottorato, Università di Firenze, Italy.

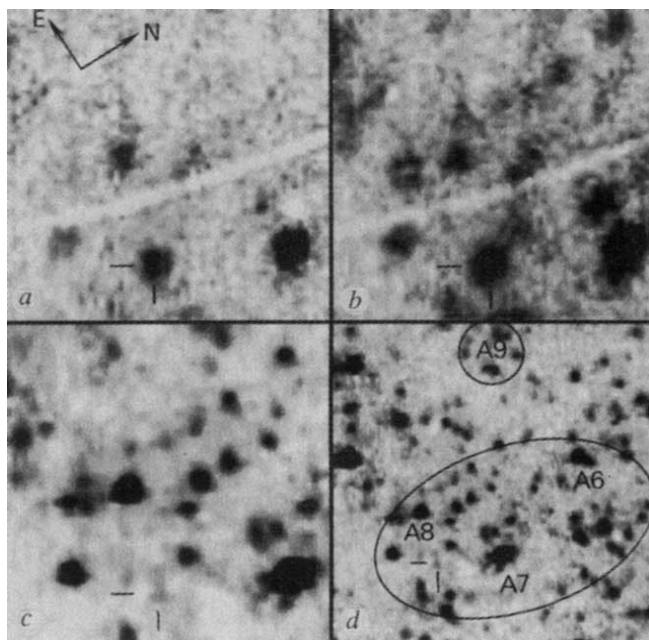


FIG. 1 Negative images of the field containing V1 taken with the F/48 camera of the FOC on the Hubble Space Telescope. *a*, A 38-min exposure with the F140W filter; *b*, a 18.7-min exposure with the F220W filter; *c*, a 14.8-min exposure with the F342W and F587W filters in combination. The first two images were taken on 28 November 1990 (ref. 4) and the third was taken on 20 February 1992. All three frames cover an area of size $2.8'' \times 2.8''$. The prominent white wavy lines running through *a* and *b* are artefacts caused by the readout beam flyback. *d*, Same exposure as in *c* but showing a field of $5.6'' \times 5.6''$ containing the smaller field shown in *c*. The X-ray source error ellipse and the stars detected by Aurière *et al.*³ in the region are marked on *d* for reference. The only blue straggler from ref. 4 in this region is HST-2 (star 6 (A6) in the list provided by Aurière *et al.*³). The orientation for all frames is marked in *a*. The position of V1 is indicated by a cross-hair in all frames. It is at the limit of detectability in the U band.

configuration was calibrated absolutely by comparison with the published fluxes of several bright objects in the same field³. This gives an excellent match between the relative ground-based fluxes and the FOC count rates of the objects chosen and provides estimated absolute accuracies of ± 0.2 mag for the entire field. No object in the field exceeds the limiting count rate for strictly linear response of the detector.

In Fig. 2 we show the average measured flux of V1 at the three passbands described above corrected for the canonical colour excess $E(B - V) = 0.04$ mag reddening for 47 Tuc¹⁷ as a function of wavelength. The two measurements at 1,650 Å and 2,300 Å were obtained with wide-band filters that are in principle affected by red leaks. In the case of V1, however, this effect is negligible because no known stellar spectrum is steep enough in the 3,500–6,000 Å region to yield the appropriate ultraviolet flux. This is the only object in the region of the core of 47 Tuc observed so far with the instrumental magnitude at 2,200 Å, m_{220} (ref. 4) greater than 20 whose spectrum is consistent with a negative continuum power-law index in the far ultraviolet. The blue stragglers are all consistent with a positive power-law index in the ultraviolet, as expected from their A–F spectral types obtained from their position on the ultraviolet colour magnitude diagram (CMD)⁴. At a distance of 4.6 kpc (ref. 17), the fluxes shown in Fig. 2 correspond to an ultraviolet optical luminosity of $\sim 8 \times 10^{32}$ erg s⁻¹. Extrapolating the black-body curve shown in Fig. 2 from the corrected $U = 19.4$ magnitude to visible wavelengths, we obtain $B = 20.8$ and $V = 21$. For a distance modulus of 13.4 (ref. 18), this corresponds to $M_V = 7.6$.

The possible temporal variability of the source was examined by analysing separately several frames taken of this field at different times. Four good images for each of the filters F140W and F220W, together with five images with the filter F342W, were available where V1 was well exposed. On each frame, we measured the magnitudes of V1 along with those of four stars surrounding it, as control objects. Whereas the flux of V1 varied considerably during this time interval, the fluxes of the control stars remained constant in time within the uncertainty range of ± 0.1 mag. The light curve obtained this way spanning an 8-hour period, is shown in Fig. 3. V1 is clearly variable in the ultraviolet over this time span with an amplitude of about 1.2 mag in both bands and with a possible periodicity of ~ 6 hours. Variability in U could not be assessed reliably because of the faintness of the object in this band and the short time span over which these

latter observations were taken. Nevertheless, these images showed the U magnitude varying by at least 1 mag over 2 hours, a result which is consistent with the variability found in the ultraviolet.

Aurière *et al.*³ proposed that star 9 (A9) in their list was a possible optical counterpart of X0021.8–7221. This star is very faint in our ultraviolet and U images and resolved into a clump of at least four stars (circled in Fig. 1*d*). None of these give any indication of being variable or blue at present¹⁹.

V1 could be a faint subdwarf B star in 47 Tuc or in the foreground, a bright subdwarf in the background Small Magellanic Cloud or a background quasar. The first three possibilities are unlikely, mainly because of the short-term temporal variability but also because 47 Tuc does not have a blue horizontal branch¹⁸ because of its high metallicity, and also because the source is so close to the core. The fourth is not so easy to rule out, as the slope and level of the spectrum are consistent with this hypothesis²⁰. Although the temporal variability occurs on

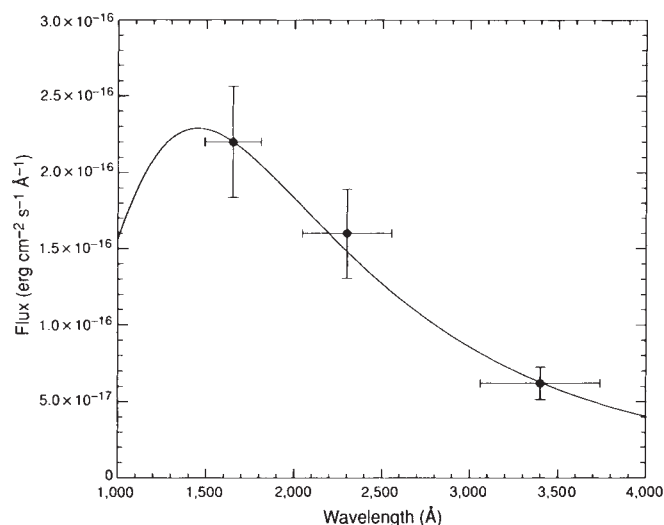


FIG. 2 Measured fluxes of V1 at the three passbands described in the text. The solid line corresponds to a black body of temperature 20,000 K normalized to the observed flux at 1,650 Å.

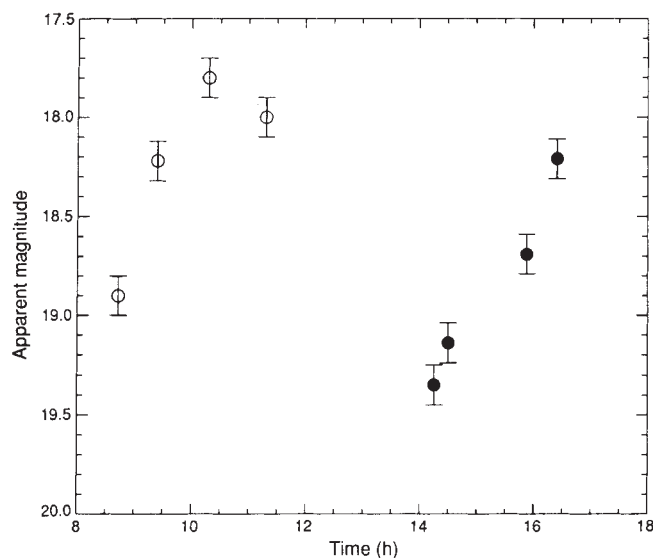


FIG. 3 The instrumental magnitude (ref. 4) of V1 as a function of time. (○) F140W data; (●) F220W data. Error bars correspond to 1σ . Time units are hours after 28 November 1990, 0:00:00 UT.

a much shorter timescale than that so far observed in a typical quasar or even a variable BL Lac object in the ultraviolet²¹, this may be due solely to the current observational limitations in this wavelength band. Variations on this and even shorter timescales are observed in the X-ray band, and the same could happen in the ultraviolet.

Our data rule out some other possibilities. The V1 spectrum is incompatible with that of a red main-sequence companion in a quiescent soft X-ray transient, or with a rotating giant or subgiant with hotspots. Its optical luminosity is too low for it to be a low mass X-ray binary at high inclination²². On the other hand, the shape and intensity of the broad-band spectrum of V1 shown in Fig. 2 are consistent with those of a typical accretion disk around a quiescent field CV²³. A normal CV, however, is expected to emit less than 10^{33} erg s⁻¹ in the 0.2–4 keV range²⁴ rather than the observed 10^{34} erg s⁻¹ of X0021.8–7221 (ref. 3). If V1 is the counterpart to X0021.8–7221, the latter's high X-ray luminosity L_X implies a ratio $L_X/L_{\text{opt}} < 0.01$, given the general correlations for normal CVs²⁴, which is incompatible with our observations. A way around this discrepancy is to assume that the source is a magnetic CV in which the strong magnetic field on the white dwarf disrupts the inner part of the accretion disk so that L_X/L_{opt} can take on the observed value of 10–20. This possibility is, in fact, consistent with all the available data on the X-ray source and V1 as its counterpart. The accretion disk spectrum, the 120-s quasi-periodicity³ of the X-ray source and the possible 6-hour periodicity in the ultraviolet flux all suggest an asynchronously rotating DQ Her type binary^{6,25,26}.

An intermediate polar binary thus gives the most natural explanation for the data, making V1 a strong candidate for the counterpart. The only remaining uncertainty concerns the higher-than-average L_X for X0021.8–7221, which is four times the luminosity of the brightest galactic CV³. It is not clear, therefore, whether V1 simply lies at the tip of the CV X-ray luminosity function or is a member of a new class of sources. The existence of such a hard binary in the core of 47 Tuc is not unexpected^{6,27,28}. If it is confirmed to be a CV, it would be the first object of this type found in the core of a globular cluster. These preliminary results clearly need to be tested by continued ultraviolet and optical observations of the 47 Tuc core coupled with higher-sensitivity measurements in the X-ray range with Rosat. Such a study could produce exciting new results on the

nature and the effect of close binaries on the evolution of dense globular clusters such as 47 Tuc. □

Received 26 May; accepted 4 September 1992.

- Hertz, P. & Grindlay, J. E. *Astrophys. J.* **275**, 105–119 (1983).
- Grindlay, J. E., Hertz, P., Steiner, J. E., Murray, G. S. & Lightman, A. P. *Astrophys. J.* **282**, L13–L16 (1984).
- Aurière, M., Koch-Miramond, L. & Ortolani, S. *Astr. Astrophys.* **214**, 113–122 (1989).
- Paresce, F. *et al. Nature* **352**, 297–301 (1991).
- Bailyn, C. D., Grindlay, J. E., Cohn, H. & Lugger, P. M. *Astrophys. J.* **331**, 303–312 (1988).
- Channugam, G., Ray, A. & Singh, K. P. *Astrophys. J.* **375**, 600–607 (1991).
- Hertz, P. & Wood, K. S. *Astrophys. J.* **290**, 171–184 (1985).
- Verbunt, F. & Meylan, G. *Astr. Astrophys.* **203**, 297–305 (1988).
- Machin, G., Callanan, P. J., Charles, P. & Naylor, T. in *Accretion-Powered Compact Binaries*, (ed. Mauche, C. W.) 163–166 (Cambridge Univ. Press, 1990).
- Bailyn, C. D., Grindlay, J. E. & Garcia, M. R. *Astrophys. J.* **357**, L35–L37 (1990).
- White, N. E. & Mason, K. O. *Space Sci. Rev.* **40**, 167–194 (1985).
- Verbunt, F., van Paradijs, J. & Elson, R. *Mon. Not. R. astr. Soc.* **210**, 899–914 (1984).
- Pallavicini, R. *Astr. Astrophys. Rev.* **1**, 177–207 (1987).
- Goodman, J. in *Dynamics of Dense Stellar Systems* (ed. Merritt, D.) 183–193 (Cambridge Univ. Press, 1989).
- De Marchi, G., Paresce, F. & Ferraro, F. *Astrophys. J.* (in the press).
- Paresce, F. *Faint Object Camera Instrument Handbook, Version 3.0* (ST ScI, Baltimore, 1992).
- Webbink, R. F. in *Dynamics of Star Clusters* (eds Goodman, J. & Hut, P.) 541–577 (Cambridge Univ. Press, 1985).
- Hesser, J. E. *et al. Proc. astr. Soc. Pacif.* **99**, 739–808 (1987).
- Bailyn, C. D. in *Accretion-Powered Compact Binaries* (ed. Mauche, C. W.) 55–58 (Cambridge Univ. Press, 1990).
- Francis, P. J. *et al. Astrophys. J.* **373**, 465–470 (1991).
- Urry, C. M., Kondo, Y., Hackney, K. R. H. & Hackney, R. L. *Astrophys. J.* **330**, 791–802 (1988).
- Verbunt, F. *Astr. Astrophys. Suppl.* **71**, 339–361 (1987).
- Van Paradijs, J. in *Accretion Driven Stellar X-ray Sources* (eds Lewin, W. H. G. & Van den Heuvel, E. P. J.) 189–260 (Cambridge Univ. Press, 1983).
- Patterson, J. & Raymond, J. *Astrophys. J.* **292**, 535–549 (1985).
- Frank, J. in *The Evolution of Galactic X-ray Binaries*, (eds Truemper, J., Lewin, W. H. G. & Brinkmann, W.) 295–300 (Reidel, Dordrecht, 1986).
- Mauche, C. W., Miller, G. S., Raymond, J. C. & Lamb, F. K. in *Accretion-Powered Compact Binaries* (ed. Mauche, C. W.) 195–202 (Cambridge Univ. Press, 1990).
- Meylan, G., Dubath, P. & Mayor, M. *Astrophys. J.* **383**, 587–593 (1991).
- Manchester, R. N. *et al. Nature* **352**, 219–221 (1991).

ACKNOWLEDGEMENTS. We thank A. Nota, W. Hack, P. Hodge, R. Jedrzejewski, P. Greenfield, D. Baxter and P. Stanley for support and assistance. G.D.M. acknowledges partial support from the ST Science Institute director's discretionary research fund. F.R.F. acknowledges the hospitality of the ST Science Institute, where he carried out part of this work.

White dwarfs as quantum crystals

Gilles Chabrier*, N. W. Ashcroft† & H. E. DeWitt‡

* Laboratoire de Physique, Ecole Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France

† Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501, USA

‡ Lawrence Livermore National Laboratory, Livermore, California 94550, USA

WHITE dwarfs are the most common endpoint of stellar evolution. Thermonuclear reactions have ceased, and the star settles into a compact object governed by relativistic, almost degenerate electrons. Bare nuclei of carbon and other heavier elements, normally considered as classical particles moving in the degenerate electron sea, are forced into a crystal lattice in the interior¹ as the star cools, and a freezing front moves outward. Numerical simulations of the freezing of classical point charges in a uniform neutralizing background have been used to model the crystallization of white dwarfs^{2–4}, but we show here that the classical approximation is not a good one: the energy per particle is significantly modified by quantum effects. The freezing process is then a transformation of a quantum liquid to a quantum solid, and the temperature of freezing may be reduced from the classical value. For lower mass stars particularly, the quantum corrections in liquid and solid seem to be comparable, and the physical conditions at freezing, classically typified by the ratio of the root-mean-square atomic displacement to the nearest-neighbour distance, are not greatly changed. Nevertheless, these new considerations may have an important effect on the cooling rate of white dwarfs, and thereby on their inferred evolution and ages.

Understanding the cooling processes in white dwarfs is important for the determination of the age of the galactic disk^{5,6} and

the properties of the halo⁷, as well as for the theory of formation of type I supernovae and neutron stars⁸. The mass distribution function of white dwarfs peaks strongly around $M \approx 0.6$ solar masses ($M_\odot$); their interiors are mostly carbon and oxygen. Some white dwarfs have masses from $1M_\odot$, to the upper limit given by the Chandrasekhar mass $M \approx 1.4M_\odot$ and consist of oxygen, neon and magnesium¹. The internal structure of these stars is essentially a carbon-oxygen core surrounded by a thin hydrogen-helium envelope. Typical central densities range from $\rho_c \approx 10^6 \text{ g cm}^{-3}$ for $0.6M_\odot$ to $\rho_c \approx 10^9 \text{ g cm}^{-3}$ for $1.4M_\odot$; central temperatures T_c of even the coolest white dwarfs are $10^6 - 10^7 \text{ K}$. Under these conditions, the atoms are fully ionized and the equilibrium of the star against gravitation is provided by the pressure of the strongly degenerate electron gas, which is largely homogeneous. For this reason white dwarfs have been thought to be one of the better natural realizations of the well-known one-component plasma (OCP) model (and its extension to binary ionic mixtures, the BIM model), a classical system of point charges immersed in a uniform neutralizing electron background.

This model has proved extremely useful in the general treatment of the properties of dense ionized matter^{2,3}. Its thermodynamic properties are characterized universally by the classical coupling parameter $\Gamma = (Z^2 e^2) / a_i kT$ (where $a_i = (4\pi/3 N_i/V)^{-1/3}$ is the mean interionic distance, for N_i nuclei in a volume V , and Ze is the nuclear charge), which measures the ratio of the classical average potential and kinetic energies. Numerical simulations by the Monte Carlo method show the OCP undergoes a fluid-solid⁴ phase transition at $\Gamma \approx 180$, which is a typical value of the coupling parameter values found in white dwarfs. Crystallization must therefore occur in white dwarfs in their final stages of evolution. The importance of strong Coulomb correlations in these stars has been emphasized^{9,10}, as has the effect of crystallization on their cooling timescales^{11,12}. The behaviour of the phase diagram and its role in the evolution of white dwarfs has been studied extensively⁷.

Crystallization has always been assumed to be nearly classical: quantum effects have been assumed small, and if treated at all they appear in the free energy of the fluid phase from an expansion to the second order of the quantum partition function in powers of Planck's constant (the Wigner-Kirkwood expansion). In this picture crystallization is reached whenever $\Gamma = \Gamma_m \approx 180$, or, put another way, whenever the free energy of the solid equals the nearly classical free energy of the fluid, which gives a similar condition^{12,13}. Here we show that quantum effects cannot be considered to be a minor perturbation either in the solid or in the liquid phase; the Wigner-Kirkwood expansion parameter is actually too large under white dwarf freezing conditions for the expansion to be useful. Proper inclusion of quantum effects can appreciably alter the freezing conditions in the more dense stars. A second point is that quantum mechanics alters the freezing condition itself, and in the low-mass stars does so in a way that leaves the classical prediction almost intact.

We first define some dimensionless characteristic parameters. The density parameter for the electron gas $r_s = a_e/a_0$ measures the mean interelectron distance in units of the Bohr radius. Similarly $R_s = a_i/a_{oi}$ measures the mean internuclear distance in units of the ionic Bohr radius ($a_{oi} = \hbar^2/M(Ze)^2$). In these units, white dwarf interiors obey $10^{-3} \leq r_s \leq 10^{-2}$. The ionic de Broglie wavelength $\Lambda = (2\pi\hbar^2/Mk_B T)^{1/2}$ is, in units of a_i , $\Lambda/a_i = \sqrt{(2\pi)(\Gamma/r_s)^{1/2}}$. The ionic plasma frequency multiplied by $\hbar$ is $\hbar\omega_p = (4\pi(Ze)^2 N_i/MV)^{1/2}$, and in units of $k_B T$ it is: $\hbar\omega_p/k_B T = \eta = \sqrt{3\Gamma/R_s^{1/2}}$. For white dwarfs at crystallization, $\Lambda/a_i < 1$ but $\eta > 1$. To get the first important dynamic energy, we note that the lattice zero point energy is

$$\frac{E_0}{N} = \frac{1}{N} \sum_{\lambda, \mathbf{q}} \frac{1}{2} \hbar \omega_{\lambda}(\mathbf{q}) = \frac{3}{2} \hbar \omega_p \mu_1 \quad (1)$$

where the $\omega_{\lambda}(\mathbf{q})$ are the phonon angular frequencies for the three polarization modes (T_1, T_2, L), of a lattice assumed to have a simple nucleus in each unit cell, and μ_1 is the first moment of the frequency spectrum (the n th moment is defined as $\mu_n = 1/3N \sum_{\lambda, \mathbf{q}} (\omega_{\lambda}(\mathbf{q})/\omega_p)^n$). Accurate Monte Carlo simulations^{14,15} give the first moment as $\mu_1 = 0.511$. The ratio of the zero-point energy to a typical classical kinetic energy is then

$$\frac{E_0}{Nk_B T} = \frac{3}{2} \eta \mu_1 \quad (2)$$

When the r.m.s. displacements of nuclei vibrating around the lattice sites are determined more by the quantum zero-point energy than by classical considerations the crystal may be referred to as a quantum crystal.

Under crystallization conditions, equation (2) yields $E_0/k_B T \approx 2$ at $r_s = 10^{-2}$ and $\sim 5-8$ at $r_s = 10^{-3}$; in other words, the zero-point energy can be several times the classical thermal energy at melting. This suggests that white dwarfs must crystallize from quantum liquids into quantum crystals, if they crystallize at all. For the thermodynamic conditions characteristic of other cool, dense astrophysical objects such as giant planets or brown dwarf stars, made up essentially of hydrogen-helium mixtures, the same analysis yields $\eta \ll 1$, implying that these objects are essentially classical.

To take our investigation further, we calculate the free energy of the crystal in the harmonic approximation using an Einstein model (single vibrational frequency) for the longitudinal mode and a Debye model (linear dispersion approximation, $\omega = ck$ where c is the speed of light and k the wavevector, for the vibrational spectrum) for the transverse modes¹⁶. The choice of parameters in the model is constrained by the accurately determined first moments¹⁵. The accuracy of the harmonic approximation for the OCP has been assessed by Monte Carlo simulations, and the representation given by the Debye and Einstein models¹⁵ is reasonable. Within these approximations, the frequencies of the two degenerate acoustic modes and the longitudinal mode are given by

$$\omega_{T_1, T_2}(\mathbf{q}) = \frac{\alpha \omega_p}{k_D} \mathbf{q}; \quad \omega_L(\mathbf{q}) = \gamma \omega_p \quad (3)$$

where $k_D = (6\pi^2 N_i/V)^{1/3}$ is the Debye wavenumber. The parameter α is first taken from the classical limit of the Lindemann parameter (see below), which yields $\alpha = \sqrt{2/\mu_2} \approx 0.393$; γ is determined from Kohn's sum rule, $\sum_{\lambda=1}^3 \omega_{\lambda}^2(k) = \omega_p^2 = 4\pi(Ze)^2 n_i/m_i$, at $q = k_D$, giving $\gamma = (1 - 2\alpha^2)^{1/2} \approx 0.832$. Under the conditions of interest, screening of the longitudinal mode by the free electrons is restricted to a negligible fraction of the Brillouin zone (the symmetric unit cell of the reciprocal lattice)¹⁶ because $k^2(\mathbf{q})/k_{TF}^2 > 1$. Here $\mathbf{q}$ is the reciprocal lattice vector corresponding to the longitudinal phonon mode with the longest wavelength¹⁷, and k_{TF} is the relativistic Thomas-Fermi wavevector. Straightforward calculations then lead to a thermal free energy per particle

$$\begin{aligned} \frac{F^{\text{th}}}{Nk_B T} &= \frac{1}{N} \sum_{\mathbf{q}, \lambda} \ln [2 \sinh \frac{1}{2} \beta \hbar \omega_{\lambda}(\mathbf{q})] \\ &= \frac{E_0}{Nk_B T} + [-\frac{2}{3} D_3(\alpha \eta) + 2 \ln(1 - \exp(-\alpha \eta)) \\ &\quad + \ln(1 - \exp(-\gamma \eta))] \end{aligned} \quad (4)$$

where $D_n(x) = n/x^n \int_0^x t^n dt/(e^t - 1)$ is the Debye integral of order n . When $\eta \ll 1$, equation (4) reduces to the standard classical thermal energy¹⁴ augmented by correction terms in $\hbar^4$ (ref. 18).

$$\begin{aligned} \frac{F_{\text{cl}}^{\text{th}}}{Nk_B T} &= \frac{3}{2} \ln \Gamma - 1.8856 + \frac{3}{2} \ln(k_B T / (Z^2 e^2 / 2a_{io})) \\ &\quad + [\frac{1}{24} \eta^2 - f(\Gamma) \eta^4] \end{aligned} \quad (5)$$

for the body-centred-cubic structure. The Wigner-Kirkwood expansion implicit in equation (5) differs between fluid and solid only at order η^4 and for either phase is valid only when quantum effects represent a minor correction to the classical thermal energy.

At melting, the liquid and the solid electrostatic (Madelung) energies exactly cancel out, so that the thermal free energy of the solid (equation (5)) must equal the ionic thermal free energy of the fluid, augmented by its perfect gas contribution. Comparison of equation (4) with equation (5) for $\Gamma = \Gamma_m$ and under white dwarf conditions shows that quantum effects increase the nearly classical free energy by ~ 10 –25%. Accordingly a simple $\hbar^2$ expansion is not accurate enough when considering white dwarf crystallization. At present, no simple free energy model has been derived for a quantum fluid at finite temperature. The significant difference between liquids and solids is the absence of shear forces over macroscopic distances in the liquid, so that transverse phonons at long wavelengths are certainly not present in the liquid state. Over short distances ($k \geq k_D/3$), however, the description of the liquid near melting may be similar to that of the solid, and if so we can expect the classical free energy of the liquid state also to be strongly influenced by quantum effects. A question arises: even though quantum effects significantly modify the classical free energies, will they substantially affect the melting curve; or will quantum contributions to the liquid and the solid cancel out? Without exact calculations, we must rely on an empirical Lindemann criterion to address this question.

In the same harmonic approximation the mean square displacement of nuclei is given by

$$\langle u^2 \rangle = (\hbar/2M) \sum_{\lambda, q} \omega_{\lambda}^{-1}(\mathbf{q}) \coth(\beta \hbar \omega_{\lambda}(\mathbf{q})/2) \\ = \left(\frac{3 \hbar \mu_{-1}}{2 M \omega_p} \right) \left\{ 1 + \frac{2 \mu_{-2}}{\eta \mu_{-1}} D_1(\alpha \eta) \right\} \quad (6)$$

or

$$\langle u^2 \rangle / d^2 = \langle u_o^2 \rangle / d^2 + (\langle u_d^2 \rangle / d^2) D_1(\alpha \hbar \omega_p / k_B T) \quad (7)$$

where $d = (3\pi^2)^{1/6} a_i$ is the nearest-neighbour distance in a body-centred-cubic lattice, and D_1 is a Debye integral. The first term in equation (7) is $\langle u_o^2 \rangle / d^2 = 0.280 \mu_{-1} / R_s^{1/2}$ and is the quantum ($T \rightarrow 0$) value of the mean square displacement normalized to the square of the nearest-neighbour distance. Correspondingly, the coefficient $\langle u_d^2 \rangle / d^2$ in the second term is the classical value ($\hbar \rightarrow 0$) and may be written $\langle u_d^2 \rangle / d^2 = 0.323 \mu_{-2} / \Gamma$. At melting these quantities become the squares of the Lindemann parameters (γ_o^2 and γ_d^2 respectively) and are known for both quantum and classical conditions. Thus $\gamma_d = 0.153$ for the classical OCP ($\Gamma_m = 180$) and $\gamma_o = 0.249$ for the quantum case (charged bosons¹⁹). Both 'experimental' values are well reproduced by equation (7) at melting conditions, confirming the essential validity of the harmonic approximation. By applying equation (7) to white dwarfs with the appropriate density at the classical temperature of crystallization (as given by $\Gamma = 180$), we find $(\langle u^2 \rangle / d^2)^{1/2} = 0.156$ and $(\langle u^2 \rangle / d^2)^{1/2} = 0.155$ at $r_s = 10^{-2}$ for pure

carbon and oxygen, respectively. At higher densities ($r_s = 10^{-3}$) we find $(\langle u^2 \rangle / d^2)^{1/2} = 0.176$ and $(\langle u^2 \rangle / d^2)^{1/2} = 0.164$. To see how these normalized r.m.s. displacements compare with the proper Lindemann melting parameters appropriate to the given conditions, we need to know how the Lindemann parameter varies between its classical value (0.153) and its quantum value (0.249) for the intermediate case. Mochkovitch and Hansen²⁰ provide the interpolation

$$\gamma = \gamma_o - a / (1 + b \eta^2) \quad (8)$$

where $a = 0.096$, $b = 0.341 \times 10^{-2}$ and, as noted above, $\eta \approx \Gamma / R_s^{1/2}$. Here the value of b corresponds to the inclusion of transverse modes; its value including the longitudinal mode is 0.51×10^{-2} (the results are not particularly sensitive to either). The interpolation (8) accurately recovers both the quantum limit ($T \rightarrow 0$) and the classical limit ($\hbar \rightarrow 0$), and also provides an accurate second-order expansion in η .

If we accept equation (8) and the Lindemann criterion for melting, then we find $\gamma(C) = 0.171$ and $\gamma(O) = 0.161$ for $r_s = 10^{-3}$ and $\gamma(C) = 0.155$ and $\gamma(O) = 0.153$ for $r_s = 10^{-2}$. It follows that at temperatures which would be sufficient to crystallize the star classically, the additional quantum effects will inhibit such crystallization and the system will remain a partially quantum fluid. At a given density a lower temperature will be required to crystallize the star. For example, for a pure carbon star, solving consistently equations (7) and (8) gives $\Gamma_m = 183$ at $r_s = 10^{-2}$ and $\Gamma_m = 199$ at $r_s = 10^{-3}$. Quantum effects lower the melting temperature by $\sim 10\%$ for the denser stars, whereas the classical crystallization temperature remains almost unaffected for the less massive stars. Given the form of the interpolation equation (8) and the semiempirical nature of the Lindemann criterion, the results can only be viewed as estimates. Nevertheless they show that quantum effects are likely to modify appreciably the melting curve of massive white dwarfs ($M > 1 M_\odot$). These stars represent a substantial portion of the white dwarf population, especially in binary systems, and are of particular importance for the formation of type I supernovae. The luminosity of cool white dwarfs scales as $L_m \sim T_c^{5/2}$ (refs 20, 21) whereas the time delay produced by crystallization is $\Delta \tau \sim E / L_m(T)$ where E is the gravitational energy and latent heat released by crystallization. Therefore, a 10% decrease in the temperature at crystallization yields a 25% decrease in the luminosity, and significantly shortens the lifetime of the star both before and after crystallization. The other main result of our calculations is that for most white dwarfs, quantum effects are of comparable importance in the fluid and in the solid phase. This suggests that white dwarfs are already in a notably quantum state before crystallization occurs. Therefore Debye cooling is likely to occur before crystallization, and may modify the cooling history of the star. If this proves to be correct, standard white dwarf cooling theory should be revised to include a proper treatment of quantum effects. Given the key role of the luminosity function in the determination of the age of the galactic disk^{5-7,22,23}, a better treatment of these will be needed to determine the age of a white dwarf for a given luminosity. $\square$

Received 12 May; accepted 9 September 1992.

- Shapiro, S. L. & Teukolsky, S. A. *Black Holes, White Dwarfs, and Neutron Stars*, Ch. 3 & 4, (Wiley, New York, 1983).
- Brush, S. G., Sahlin, H. L. & Teller, E. *J. chem. Phys.* **45**, 2102 (1966).
- Hansen, J. P. *Phys. Rev. A* **8**, 3096 (1973).
- Stringfellow, G. S., De Witt, H. E. & Slattery, W. I. *Phys. Rev. A* **41**, 1105 (1990).
- Winget, D. E. *et al. Astrophys. J.* **315**, L77 (1987).
- Garcia-Berro, E., Hernanz, M., Mochkovitch, R. & Isern, J. *Astr. Astrophys.* **193**, 141 (1988).
- Mochkovitch, R., Garcia-Berro, E., Hernanz, M., Isern, J. & Panis, J. F. *Astr. Astrophys.* **233**, 456 (1990).
- Canal, R., Isern, J. & Labay, J. *Nature* **296**, L225 (1982); *A. Rev. Astr. Astrophys.* **28**, 183 (1990).
- Abrikosov, A. *JETP* **12**, 1254 (1961).
- Salpeter, E. E. *Astrophys. J.* **134**, 669 (1961).
- Mestel, L. & Ruderman, M. A. *Mon. Not. R. astr. Soc.* **136**, 27 (1967).
- Van Horn, H. M. *Astrophys. J.* **151**, 227 (1968).
- Lamb, D. Q. & Van Horn, H. M. *Astrophys. J.* **200**, 306 (1975).

- Hansen, J. P. & Pollock, E. L. *Phys. Rev. A* **8**, 3110 (1973).
- Albers, R. C. & Gubernatis, J. E. *Phys. Rev. B* **23**, 2782 (1981); **33**, 5180 (1986).
- Ashcroft, N. W. & Mermin, N. D. *Solid State Physics* (Saunders, Philadelphia, 1976).
- Kugler, A. A. *Ann. Phys.* **53**, 133 (1969).
- Hansen, J. P. & Viellefosse, P. *Phys. Lett. A* **53**, 187 (1975).
- Ceperley, D. & Alder, B. *Phys. Rev. Lett.* **45**, 566 (1980).
- Mochkovitch, R. & Hansen, J. P. *Phys. Lett. A* **73**, 35 (1979).
- Nagara, H., Nagata, Y. & Nakamura, T. *Phys. Rev. A* **36**, 1859 (1987).
- Wood, M. A. & Winget, D. E. in *White Dwarfs* (ed. Wegner, G.) 282 (Springer, Berlin, 1989).
- Wood, M. A. thesis, Univ. of Austin (1990); *Astrophys. J.* (in the press).

ACKNOWLEDGEMENTS. We thank R. Mochkovitch and H. M. Van Horn for discussion. This work was initiated at the Institute for Theoretical Physics of the University of California at Santa Barbara which is supported in part by the NSF. N.W.A. is also supported by NSF at Cornell University. The work of H.E.D. was done under the auspices of the US Department of Energy at the Lawrence Livermore National Laboratory.

Real-time observation of vortex lattices in a superconductor by electron microscopy

K. Harada, T. Matsuda, J. Bonevich, M. Igarashi, S. Kondo, G. Pozzi*, U. Kawabe† & A. Tonomura

Advanced Research Laboratory, Hitachi, Ltd, Hatoyama, Saitama 350-03, Japan

* Department of Material Science, University of Lecce, via Arnesano, 73100 Lecce, Italy

† Central Research Laboratory, Hitachi, Ltd, Kokubunji, Tokyo 185, Japan

THE dynamic behaviour of the quantized vortices of magnetic flux that penetrate type II superconductors, and specifically their interaction with pinning sites, is a topic of both scientific and technological interest, particularly since the discovery of high-transition-temperature (high- T_c) superconductors¹. Until now, however, it has not been possible to study this behaviour in 'real time'. The Bitter technique², scanning tunnelling microscopy³ and scanning electron microscopy⁴ have so far provided only static images, whereas a magneto-optical technique that provides time-resolved information⁵ does not resolve individual vortices. Here we report the real-time observation of vortices in a thin film of niobium, using the technique of Lorentz microscopy⁶. The coherent and penetrating beam of a recently developed 300-kV field-emission electron microscope⁶ allows us to observe, at 30 frames per second, the motion of thermally activated vortices, and their response to an applied magnetic field.

Lorentz microscopy⁶ has been used to observe magnetic domain structures in ferromagnetic thin films. The domain walls, unseen in an in-focus electron micrograph, become visible as black or white lines upon defocusing. Two electron beams transmitted through neighbouring domains are deflected in different directions so that the defocusing makes the two beams overlap or separate, thus increasing or decreasing the intensity along the domain wall. If the magnetic flux to be observed is less than h/e (where h is Planck's constant) as in the present experiment, then this description in terms of geometric optics must be replaced by a wave-optical one⁸ taking into account the quantum-mechanical phase shifts or the rotation of the wavefront of the electron waves caused by the magnetic vector potential (compare with the Aharonov-Bohm effect⁹).

Previous attempts¹⁰⁻¹⁵ to apply this method to vortex observation have had no success, because the wavefront of an electron passing through a single vortex is rotated through a very small angle ($\sim 10^{-6}$ rad). Our electron microscope⁷ produces a coherent electron beam with a divergence angle much less than 10^{-6} rad and has high energy, 300 keV, sufficient to penetrate the sample without losing the beam collimation.

Our experimental arrangement is shown in Fig. 1. A thin film of niobium ($T_c = 9.2$ K, resistance ratio $R(300\text{ K})/R(10\text{ K}) \approx 20$) for transmission observation was prepared by chemically etching a piece $2 \times 2\text{ mm}^2$ by $30\text{ }\mu\text{m}$ thick cut from a rolled Nb film, which had been annealed at $2,200\text{ }^\circ\text{C}$ in a vacuum of 1×10^{-8} torr to increase its grain size up to $200\text{--}300\text{ }\mu\text{m}$. The film had a [110] surface and one or two holes in the middle, its thickness tapering off towards the hole; we observed the area near the hole edge which was $700 \pm 200\text{ }\text{\AA}$ thick.

The sample was positioned on a low-temperature stage, and tilted at 45° to the vertically incident electron beam, so that the electrons could sense the vortex magnetic fields penetrating the sample perpendicularly to its surface; the penetrating field was due to an external magnetic field applied horizontally. The objective lens was turned off to remove its vertical magnetic field, and the intermediate lens was used instead for focusing.

Wave-optical calculation¹⁵ shows that the vortices should be visible in an appropriately defocused image (the Lorentz micro-

graph) as a tiny globule, one side bright and the other side dark with the dividing line parallel to the projection of the vortex field on the image plane. In the present case, the optimum defocusing distance was 10 mm .

Increasing the applied magnetic field B . The sample was first cooled to 4.5 K , and B was then gradually increased. A few vortices appeared at $B = 32\text{ G}$, which gives the lower critical field H_{c1} of our film sample. When B was increased further, the number of vortices increased almost instantly, as expected from electron microscopy indicating that the sample was free from strong pinning sites such as crystal dislocations, voids and grain boundaries.

Some vortices (10–20% of those seen) continued to move for a few minutes before calming down, the length of relaxation time depending on the region of the film observed. They fluctuated around their own sites with amplitudes of $\sim 0.2\text{ }\mu\text{m}$ and frequencies of $\sim 1\text{ s}^{-1}$, occasionally hopping a distance of $\sim 0.5\text{ }\mu\text{m}$. These weak pinning sites are thought to be interstitial or substitutional atoms. The vortices increased in number with B , to become hexagonally arranged at $B \approx 100\text{ G}$.

An example of the equilibrium Lorentz micrographs at $B = 100\text{ G}$ is shown in Fig. 2. The sample film has a fairly uniform thickness in the region shown, but is bent along the black curves ('bend contours') which are due to Bragg reflections at atomic planes brought to a favourable angle by bending. Each globule with black and white contrast is the image of a vortex (see Fig. 1) as we explained before. We also confirmed by measuring the electron phase shift due to a vortex that the image corresponded to a single vortex carrying one flux quantum $h/2e$. This contrast was reversed as it should when B was reversed. The density of the vortices is about $5\text{ }\mu\text{m}^{-2}$ as expected for $B = 100\text{ G}$; in fact $(100\text{ G})/(h/2e) = 4.8\text{ }\mu\text{m}^{-2}$. As the black (or white) sides of all

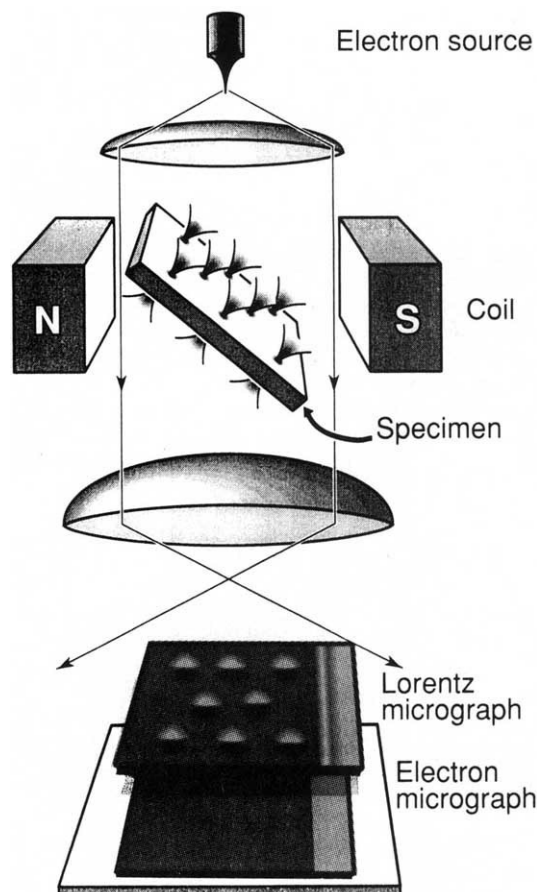
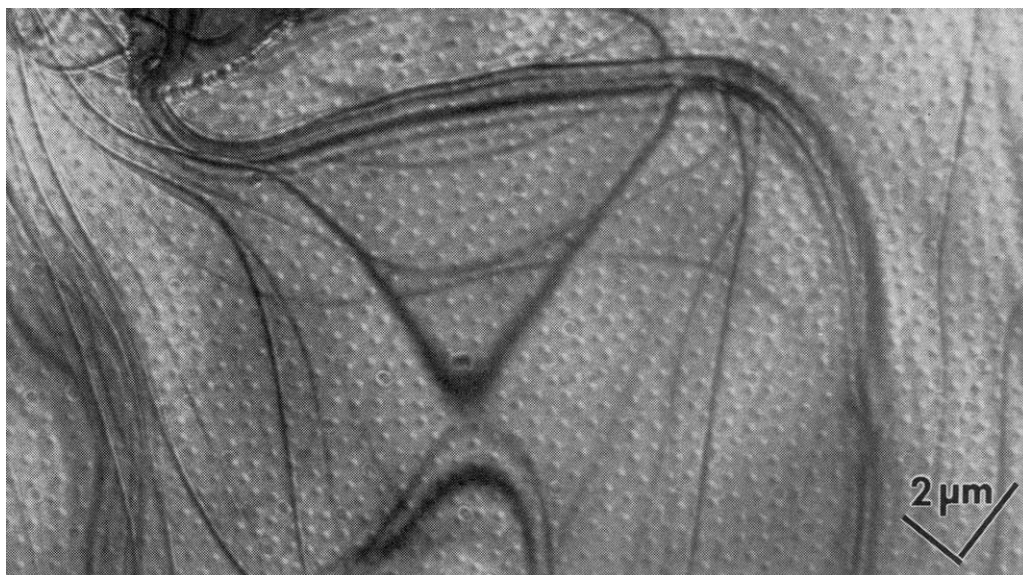


FIG. 1 Schematic diagram of vortex-lattice observation.

FIG. 2 Lorentz micrograph of a Nb thin film. A magnetic field of 100 G is applied to the film in the superconducting state at 4.5 K. Vortices are observed as small globules, each with black and white contrast. The magnification differs in the two directions by $\sqrt{2}$, because the film is tilted by 45° . The dark lines are bend contours.



the globules have the same orientation, the polarities of all the vortices in the region are the same.

When B was increased to 150 G, the vortex lattice was compressed to a more regular hexagonal lattice with a lattice spacing of $0.35 \mu\text{m}$. Apparently the vortices were influenced in configuration as well as in movements by weak pinning sites in the Nb sample.

Raising the sample temperature T . As T was raised gradually from 4.5 K to 8 K at fixed magnetic field $B = 150$ G, vortices were released from pinning sites and hopped to form a more and more regular hexagonal lattice, as shown in Fig. 3. The vortex diameters increased with T because of the increase in the penetration depth, and eventually the vortices disappeared at $T_c = 9.2$ K.

When the temperature of the same sample was gradually lowered from 10 K to the initial 4.5 K, the vortex lattice, although disturbed by hopping, was more regular than it had been initially, showing that an irreversible transformation had occurred.

Removal of the applied field B . It took 6 s for the applied field to decay to zero from $B = 150$ G when the coil current was switched off. During this period the electron beam was deflected out of the field of view. When it returned, 90% of the vortices were gone. Those which remained, being trapped at pinning sites, slowly moved towards the hole in the sample, where they disappeared. They moved by hopping in the direction of decreasing vortex density, seemingly along preset routes according to local variations in thickness. Besides these, there were some that

oscillated repeatedly around fixed centres, with amplitudes $\sim 0.3 \mu\text{m}$ and frequencies $\sim 10 \text{ s}^{-1}$. Examples of the vortex movements as reproduced from our videotape are shown in Fig. 4. Figure 4a and b shows the instants when a vortex happened to start and stop hopping, respectively, as indicated by the arrows; the interval of time is 0.1 s. In the succeeding 1.3 s between b and c, three vortices hopped, as indicated by the arrows. That each arrow in c stands for a single hop was checked by examining all the frames in between.

We conclude that Lorentz microscopy with a 300-kV field-emission electron microscope opens up new opportunities for real-time observation of the microscopic dynamics of individual vortices. We can observe not only flux creep, pinning and so on in motion, but also tell polarities and flux values of individual vortices. The spatial resolution of the technique is a few hundred ångströms in this experiment, but can be improved to a few ångströms by adjusting the image defocusing. Time resolution is $1/30$ s at present, and may be improved to $1/300$ s. Magnetic flux can be resolved down to $h/6e$, one-third of the elementary flux. The applied magnetic field could be varied up to 150 G in our experimental arrangement with its horizontal field, but the upper limit can be 10 kG if the (vertical) field of the objective lens is used.

We hope to apply our method to observe flux melting and other novel features of vortices in high- T_c superconductors. There are some technical problems to solve. First, the sample must be a self-supporting thin film with thickness of $\sim 1,000 \text{ Å}$ for transmission observation. Some high- T_c materials can be

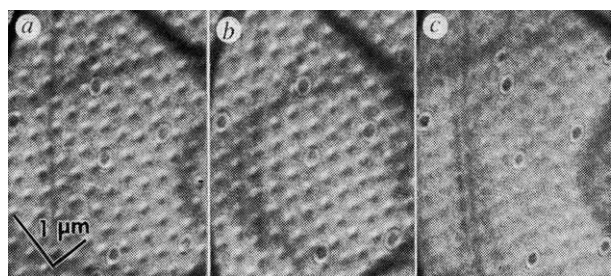


FIG. 3 Dependence of the vortex images on temperature. a, $T = 4.5$ K; b, $T = 6.0$ K; c, $T = 8.0$ K. The vortices become larger as the penetration depth increases with temperature. The vortex configuration, random at 4.5 K, turns into a more and more regular hexagonal lattice at higher T , which can be recognized more easily when the photographs are seen obliquely.

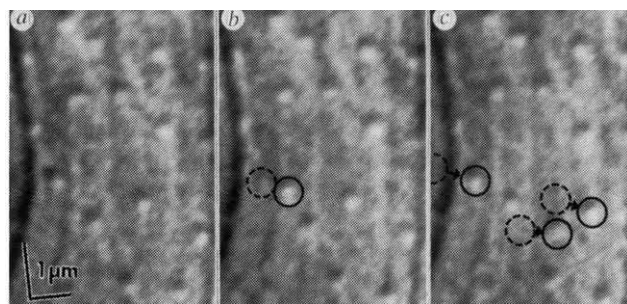


FIG. 4 Dynamics of the vortices as seen at times t after magnetic field has been switched off from 150 G. a, $t = t_0 + 0$ s; b, $t = t_0 + 0.1$ s; c, $t = t_0 + 1.4$ s ($t_0 = 170$ s). Dotted and solid circles show vortex positions before and after hopping, and the hopping directions are indicated by the arrows.

grown only on a single-crystalline substrate, and its removal could destroy them. Second, the sample must be placed in a high vacuum, where it may suffer oxygen desorption, to which the high- T_c materials are very sensitive. Third, the larger penetration depths in such materials may require substantial changes in the observation conditions. As Lorentz microscopy relies on

defocusing, its spatial resolution is necessarily less than that of electron holography^{16,17}, which, however, is not suitable for time-resolved, real-time observations. These two techniques are thus complementary. Experiments using electron holography are now in progress to study the interaction of the vortices with small defects. □

Received 14 August; accepted 30 September 1992.

1. Bishop, D. J., Gammel, P. L., Huse, D. A. & Murray, C. A. *Science* **255**, 165–172 (1992).
2. Essman, V. & Träuble, H. *Phys. Lett. A* **24**, 526–527 (1967).
3. Hess, H. F., Robinson, R. B., Dynes, R. C., Valles, J. M. Jr & Waszczak, J. V. *Phys. Rev. Lett.* **69**, 214–216 (1989).
4. Mannhart, J., Bosch, J., Gross, R. & Huebener, R. P. *Phys. Rev. B* **35**, 5267–5269 (1987).
5. Durán, C. A. *et al. Nature* **357**, 474–477 (1992).
6. Hirsch, P. B., Howie, A., Nicholson, R. B., Pashley, D. W. & Whelan, M. J. *Electron Microscopy of Thin Crystals*, (Butterworths, London, 1965).
7. Kawasaki, T., Matsuda, T., Endo, J. & Tonomura, A. *Jap. J. appl. Phys.* **29**, L508–L510 (1990).
8. Wohlleben, D. J. *Phys. Lett.* **22**, 564–566 (1966).
9. Pestkin, M. & Tonomura, A. *The Aharonov-Bohm Effect, Lecture Notes in Physics*, Vol. 340 (Springer, Heidelberg, 1989).

10. Hirsch, P. B., Howie, A. & Jakubovics, J. P. *Proc. int. Conf. Electron Diffraction Crystal Defects*, Vol. 1, B-5 (Pergamon, Oxford, 1965).
11. Yoshioka, H. *J. phys. Soc. Jpn* **21**, 948–954 (1966).
12. Goringe, M. J. & Jakubovics, J. P. *Phil. Mag.* **15**, 393–403 (1967).
13. Wohlleben, D. *J. appl. Phys.* **38**, 3341–3352 (1967).
14. Colliex, C., Jouffrey, B. & Kleman, M. *Acta crystallogr.* **A24**, 692–696 (1968).
15. Capiluppi, C., Pozzi, G. & Valdrè, U. *Phil. mag.* **26**, 865–883 (1972).
16. Matsuda, T. *et al. Phys. Rev. Lett.* **62**, 2519–2522 (1989).
17. Tonomura, A. *Adv. Phys.* **41**, 59–103 (1992).

ACKNOWLEDGEMENTS. We thank H. Ezawa, T. Fujita, A. Fukuhara, C. N. Yang, Y. Murayama, F. Nagata, M. Tachiki and T. Onogi for discussions, and N. Osakabe, S. Kubota, H. Kasai, T. Yoshida, J. Endo, N. Miyamoto, H. Kajiyama and S. Matsunami for technical advice and assistance.

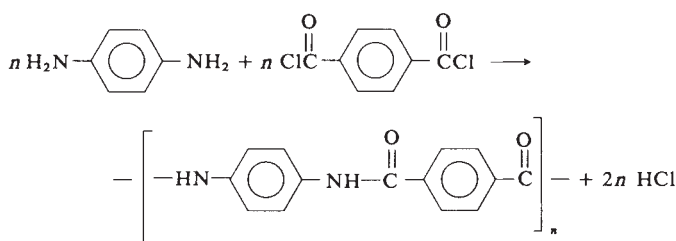
Enhancement of polymerization rates for rigid rod-like molecules by shearing

U. S. Agarwal & D. V. Khakhar

Department of Chemical Engineering, Indian Institute of Technology, Bombay, Powai, Bombay 400 076, India

RIGID, rod-like polymers are of considerable technological importance in the production of high-strength and high-modulus fibres¹. Polymerization reactions generally require near-parallel orientation of such molecules, and because of the slow rotational diffusion of the reacting species², the rate of polymerization is found to decrease significantly in the later stages of the reaction; this ultimately limits the molecular weight of the polymer. It is known, however, that rod-like molecules can be readily oriented in solution by means of imposed flow fields. Here we examine the effect of a shear flow on the polymerization reaction between *p*-phenylene diamine and terephthaloyl chloride, where the product, poly(*p*-phenylene terephthalamide), has a rigid, rod-like structure¹. We find that once the molecular weight of the reactants exceeds a critical value, shearing of the reaction mixture induces a pronounced orientation of the molecules, accompanied by a significant increase in the polymerization rate. The technological implications are clear: exposure of polymerizing mixtures of rod-like molecules to high shear rates can result in a dramatic enhancement in the molecular weight of the resulting polymer.

The polymerization is carried out according to the following reaction¹



in a two-reactor system (Fig. 1). The solvent used in the polymerization is a 3:1 mixture of *N*-methyl pyrrolidone (NMP) and hexamethyl phosphoramide (HMPA). Powdered *p*-phenylene diamine (PPD) is dissolved in the solvent to yield a 0.2 M solution, and cooled to 3 °C in reactor I. The polymerization is initiated by dissolving powdered terephthaloyl chloride (TPC) in reactor I by means of a high-speed stirrer. On addition of

TPC, the temperature of the solution increases to 22 °C in less than 10 s, and after 1 minute of mixing it falls to 10 °C. Part of the reacting mixture is then transferred by means of a peristaltic pump to reactor II. Stirring in reactor I is discontinued, and the bob in reactor II is rotated at a desired speed. In another minute, the temperatures of reactor I and II reach 3 °C and 5 °C, respectively, and the polymerization continues in both reactors. The choice of geometry of reactor II ensures a nearly homogeneous shear flow over the entire volume, and mixing in the flow is expected to be minimal³. The powdering and weighing of the monomer, and the reaction, are done in a glove box where moisture content has been reduced to less than 100 p.p.m. by drying with KOH pellets.

We study the kinetics of the reaction by quenching the reaction with water at a desired time. The precipitated polymer is dried at 80 °C under vacuum for 10 h, and the weight-average molecular weight ($\bar{M}_w$) is found by measuring the inherent viscosity of a 0.5% solution of the polymer in 98% sulphuric acid at 30 °C, and interpolating from previously reported experimental data⁴ of $\bar{M}_w$ against intrinsic viscosity. The weight-average degree of polymerization ($\bar{DP}_w$) is then obtained from $\bar{DP}_w = \bar{M}_w/238$, where 238 is the molecular weight of the repeat unit.

Figure 2a shows the evolution of the $\bar{DP}_w$ with time for reactor I and three different shear rates ($\dot{\gamma}$) in reactor II. Each data point represents a different experiment. The results indicate that for reactor I the $\bar{DP}_w$ initially increases rapidly and then levels

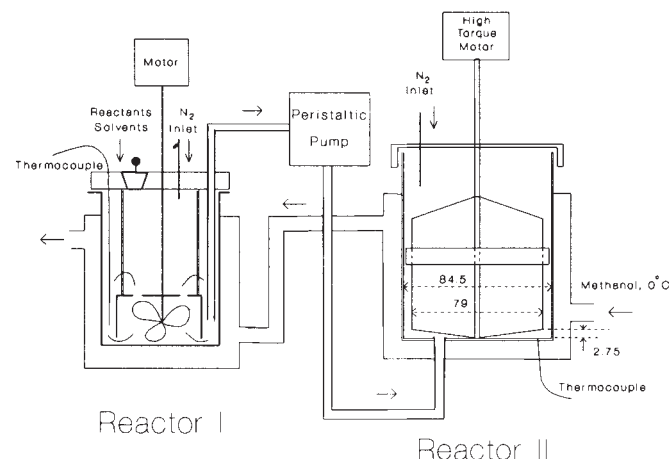


FIG. 1 Schematic view of the apparatus for polymerization. The reactor dimensions shown are in millimetres.

off, corresponding to a decrease in rate constant as reaction proceeds. This indicates diffusion-limited reaction². Behaviour in reactor II is similar initially, but is later followed by a second stage of rapid increase, depending on the shear rate. The change in $\overline{DP}_w$ with shear rate is significant (more than 60%) as compared with variation between different experimental runs ($\pm 15\%$). Figure 2b shows the $\overline{DP}_w$ at four different times for different shear rates. The flow increases the final conversion when the degree of polymerization and shear rates are sufficiently high. The applied shear rate has no significant effect until about 13 min of reaction, or $\overline{DP}_w$ up to 45.

A possible cause for the enhancement in polymerization rate could be improved mixing at the higher shear rates which would reduce the local stoichiometric imbalances of the reactants. But because the enhancement in the reaction rate becomes significant only at sufficiently high shear rates and molecular weights (Fig. 2a and b), the effect is probably not due to improved mixing. We further establish this by a set of experiments involving two different histories of shearing such that the total strain, given

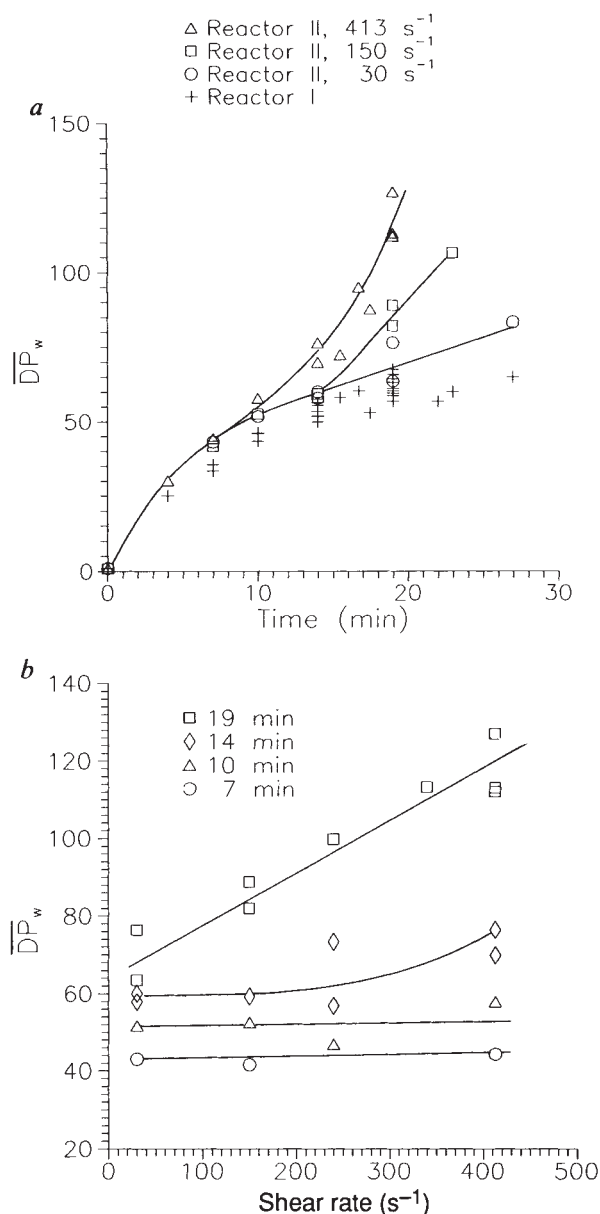


FIG. 2 a, Increase in the degree of polymerization ($\overline{DP}_w$) of the polymer in reactor I and reactor II (Fig. 1) with time, at different shear rates. b, Effect of shear rate on $\overline{DP}_w$ in reactor II (Fig. 1), for different times of reaction.

TABLE 1 Results of experiments with different shearing histories

Experiment	1-7 min	7-13 min	13-19 min	$\overline{DP}_w$
I	30 s ⁻¹	413 s ⁻¹	413 s ⁻¹	113
II	413 s ⁻¹	413 s ⁻¹	30 s ⁻¹	63

The shear rates at different times in reactor II are indicated, as well as the $\overline{DP}_w$ of PPTA obtained after 19 min of reaction.

by $\gamma = \int \dot{\gamma} dt$, is the same in both cases (Table 1). In case I, the $\overline{DP}_w$ obtained is nearly the same as the value obtained for a uniform shear rate $\dot{\gamma} = 413$ s⁻¹, whereas in case II it is close to that for $\dot{\gamma} = 30$ s⁻¹ (Fig. 2a). Thus a high shear rate beyond 13 min of reaction (case I) enhances the polymerization rate even when the extent of mixing is the same in both cases.

A second possible cause for the enhancement in polymerization rate at high $\dot{\gamma}$ could be the temperature increase of the reaction mixture resulting from viscous dissipation. Even at the largest shear rate, however, the increase was found to be only 2°C after 15 min of reaction. As the reaction rate is relatively insensitive to temperature increases (increasing the temperature from 3°C to 25°C after 13 min in reactor I gives a 5% increase in the reactor I $\overline{DP}_w$ for 27 min of reaction), viscous dissipation effects would have negligible effect in enhancing the rate.

From the experimental results, we propose the following mechanism to explain the enhancement in the polymerization rate. Initially the reaction is controlled by chemical kinetics. With increasing molecular size, the rotational diffusivity (D_r) and the translational diffusivity perpendicular to the rod axis decrease substantially⁵, leading to diffusion control and slowing of the reaction⁶. When the molecular size becomes large enough, and if the shear rate is high enough ($\dot{\gamma}/D_r > 1$, ref. 5), significant orientation of the molecules occurs locally, increasing the number of pairs of molecules with relative orientation suitable for reaction, and thus increasing the reaction rate. On-line birefringence measurements during the polymerization in a glass reactor of geometry similar to reactor II show that substantial flow induced orientation does take place (Fig. 3), and the orientation level is consistently higher at higher shear rates. The time when the maximum orientation is achieved also corresponds to the time when the polymerization rate increases sharply.

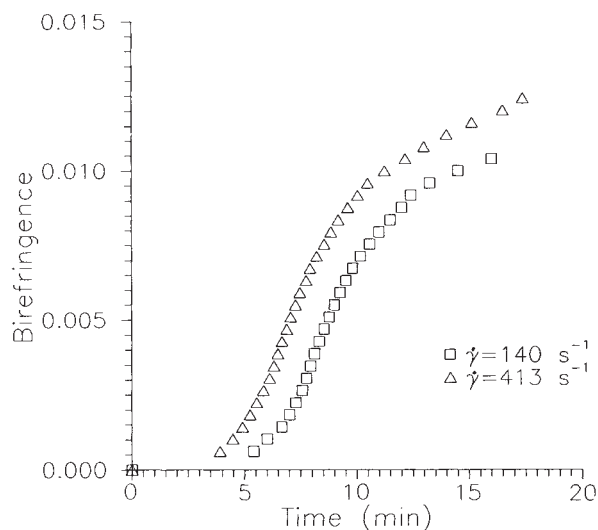


FIG. 3 Variation of birefringence (a measure of the extent of orientation of the molecules) with time for two different shear rates. The transmitted light intensity was measured using cross-polarizers for a He-Ne laser beam directed through the conical section of a glass reactor similar to reactor II. The birefringence was obtained from the peaks of the sinusoidal experimental trace of intensity against time¹¹.

An implication of the above mechanism is the tendency for longer molecules to orient to a greater degree than smaller molecules, so that there is more chance of longer molecules reacting with other long molecules, leading to a wider molecular weight distribution. The polydispersity index (p , a measure of the width of this distribution) for poly(p -phenylene terephthalamide) (PPTA) was obtained by gel permeation chromatography of samples which were made soluble in tetrahydrofuran after n -alkylation by n -octadecyl bromide^{7,8}. For 19 min of reaction and $\dot{\gamma} = 30 \text{ s}^{-1}$, the polydispersity index ($p = 2.45$) was significantly lower than that for $\dot{\gamma} = 413 \text{ s}^{-1}$ ($p = 3.25$), in agreement with the proposed mechanism.

Previous studies of polymerization of rod-like polymers have shown the necessity of high shear rates for obtaining high molecular weights^{4,9}. The type of mixer used also affects the limiting molecular weight¹⁰. Such results can be explained if the shearing causes an increase in the rate of polymerization relative to the rate of reactive end-group terminating side reactions.

The above results have relevance for systems with rotational constraints for reaction, in which slow rotational diffusion of the molecules severely reduces the reaction rate. Flow fields as

well as other orienting fields can significantly enhance such reaction rates, as well as alter the molecular weight distribution. Besides the obvious benefit of increasing production rates, the above mechanism could also be used to advantage in systems in which side reactions with small impurity molecules occur. In these cases, a better (higher-molecular-weight) product would be obtained by increasing the polymerization rate relative to the side reactions, which would be unaffected by shearing. □

Received 12 March; accepted 30 September 1992.

1. Kwolek, S. L., Morgan, P. W. & Schaefer, J. R. in *Encyclopedia of Polymer Science and Technology* (eds Mark, H. F., Bikales, N. M., Overberger, C. G. & Menges, G.) Vol. 9, 1–61 (Wiley, New York, 1987).
2. Cotts, D. B. & Berry, G. C. *Macromolecules* **14**, 930–934 (1981).
3. Chella, R. & Ottino, J. M. *Arch. Ration. Mech. Anal.* **90**, 15–42 (1985).
4. Arpin, M. & Strazielle, C. *Polymer* **18**, 591–598 (1977).
5. Doi, M. & Edwards, S. F. *The Theory of Polymer Dynamics*, 324–349 (Clarendon, Oxford, 1986).
6. Agarwal, U. S. & Khakhar, D. V. *J. chem. Phys.* **96**, 7125–7134 (1992).
7. Ogata, N., Sanui, K. & Kitayama, S. *J. Polym. Sci. Polym. Chem. edn* **22**, 865–867 (1984).
8. Takayanagi, M. & Katayose, T. *J. Polym. Sci. Polym. Chem. edn* **19**, 1138–1145 (1981).
9. Morgan, P. W. *Macromolecules* **10**, 1381–1390 (1977).
10. Vollbracht, L. in *Comprehensive Polymer Science* (eds Allen, G. & Bevington, J. C.) Vol. 5, 375–386 (Pergamon, Oxford, 1989).
11. Asada, T., Muramatsu, H., Watanabe, R. & Onogi, S. *Macromolecules* **13**, 867–871 (1980).

Cooling and freshening of the subpolar North Atlantic Ocean since the 1960s

J. F. Read & W. J. Gould

Institute of Oceanographic Sciences Deacon Laboratory, Brook Road, Wormley, Surrey GU8 5UB, UK

LITTLE is known of the interdecadal variability in the thermohaline circulation of the world's oceans, yet such knowledge is essential as a background to studies of the effects of natural and anthropogenic climate change. The subpolar North Atlantic is an area of extensive water mass modification by heat loss to the atmosphere. Lying as it does at the northern limit of the global thermohaline "conveyor belt"^{1,2}, changes in this region may ultimately have global consequences. Here we report that in August 1991 the waters between Greenland and the United Kingdom were on average 0.08°C and 0.15°C colder than in 1962 and 1981, respectively, and slightly less saline than in 1962. The cause appears to be renewed formation of intermediate water in the Labrador Sea from cooler and fresher source waters, and the spreading of this water mass from the west. Variations in the source characteristics of Labrador Sea Water can be traced across the North Atlantic, with a circulation time of 18–19 years between the Labrador Sea and Rockall Trough. More recently formed Labrador Sea Water, with even lower temperature and salinity, should cool and freshen the North Atlantic still further as it circulates around the ocean in the coming decade.

A survey of the subpolar North Atlantic was made in August 1991 as a contribution to the World Ocean Circulation Experiment. The survey (CONVEX-91) was designed to resolve the features of the gyre-scale circulation. The conductivity-temperature-depth (CTD) stations discussed here were along two roughly zonal tracks at $\sim 58^\circ\text{N}$ (north section) and 53°N (south section), linked by a section near 40°W (west section)¹ (Fig. 1). One of the most striking features of the data was the abundance and extreme characteristics of Labrador Sea Water (LSW) in the western part of the survey. LSW is one of the main water masses of the North Atlantic. It is believed to form in the central Labrador Sea by deep convection² and spreads out at intermediate depths throughout most of the area north of 40°N . It is characterized by a marked salinity minimum and has been defined³ as having potential temperature and salinity

characteristics between 3.3 – 3.4°C and 34.84 – 34.87 , respectively. The potential temperature and salinity minimum seen during CONVEX-91 was pronounced, reaching 2.83°C , 34.84 . Another distinguishing feature of LSW is its weak vertical density gradient³, and during CONVEX-91 this was found between depths of 500 and 2,000 m within the density range of $\sigma_{1.5} = 34.64$ – 34.68 kg m^{-3} (where $\sigma_{1.5}$ is potential density referenced to a depth of 1,500 db).

Changes in LSW and other water masses of the subpolar North Atlantic have already been observed over the past few years^{4–8}. The characteristics of LSW near the source show considerable variation within tens of years^{2–4,9}. The variability is believed to be caused by a combination of changes in the rate of renewal and changes in the properties of the water from which LSW is formed⁹. Monitoring over the past 50 years⁹ has shown that during the 1960s deep convection did not occur and the temperature and salinity of LSW gradually increased. Deep convection was renewed in 1971–72⁹ and resulted in an abrupt drop in temperature. (J. R. N. Lazier, manuscript in preparation) has updated the time series of ref. 3 of temperature on the density surface $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$ from which we have inferred salinity (Fig. 2a). This shows that after a freshening in the early 1970s, the salinity (and temperature) began to increase in the early 1980s, although not to the same level as in 1970. Since then there has been a further, more extreme freshening (and cooling) on this density surface. Thus there have been two

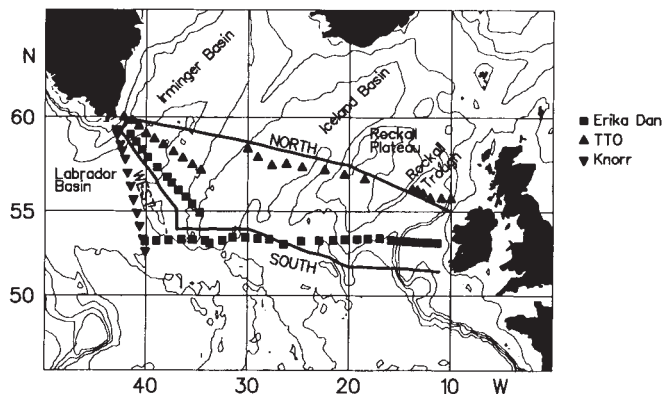


FIG. 1 Location of the CONVEX-91 sections, with Erika Dan (1962), TTO (1981), and Knorr (1983) stations, showing the topography of the region.

periods of renewal and change; during 1971–76 and again between 1984–90.

Salinity on density surfaces close to the temperature and salinity minimum of LSW from the CONVEX-91 northern section shows two main features (Fig. 2b): first, an increase in salinity from west to east (high salinity values over the Mid-Atlantic Ridge and Rockall plateau are probably the result of enhanced mixing over the topography); second, a change in the density surface of the salinity minimum.

LSW seems to be advected around the North Atlantic with modification by mixing and by time changes at the LSW source³. Salinity on the $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$ density surface at the eastern end of the north section (Fig. 2b) are higher than anything seen at the LSW source (Fig. 2a) implying that mixing has taken place. Here we will not examine the mixing processes, but accept their role in modifying the water mass as we seek to relate the time-series (Fig. 2a) to the spatial distribution (Fig. 2b). Salinity on the $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$ density surface in the Rockall Trough is ~ 34.925 , and we will start by associating this with the highest values in the Labrador Sea, ~ 34.915 , seen in 1970–71. Salinity on the same density surface in the Iceland Basin is fresher, between 34.89–34.90, so it could be older than that in the Rockall Trough, that is before 1970, or it could have originated more recently, when salinity increased between 1981 and 1985. If the circulation scheme proposed in ref. 3 is adopted, namely cyclonic flow out of the Labrador Sea eastwards across the North Atlantic, with northward branches into the Irminger Basin and into the Iceland Basin and Rockall Trough, then the water in the Iceland Basin should be of more recent origin than that of the Rockall Trough because it is closer to the source. Therefore we suggest that the LSW in the Iceland Basin formed during the period 1981–85 in the Labrador Sea.

We can corroborate our timescale of the circulation by looking at the changes in density surface of the salinity minimum. The density of LSW was about $\sigma_{1.5} = 34.66 \text{ kg m}^{-3}$ during the 1960s³, decreasing to $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$ in the early 1970s^{3,6}, and has since increased again reaching $\sigma_{1.5} = 34.66 \text{ kg m}^{-3}$ in 1986⁴ and $\sigma_{1.5} = 34.67 \text{ kg m}^{-3}$ in 1992 (J. R. N. Lazier, manuscript in

preparation). The CONVEX-91 north section (Fig. 2b) shows the salinity in the Iceland Basin to be at least on the $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$ isopycnal surface. This suggests that this water must have formed before 1986, but later than the 1960s. The salinity minimum in the Irminger Basin lies between $\sigma_{1.5} = 34.66$ and 34.67 kg m^{-3} , indicating its recent origin, and is markedly fresher than the rest of the section. The difference in both salinity and density of the salinity minimum west and east of the Mid-Atlantic Ridge suggests that more recent, denser LSW has not yet spread to the east.

The density surface of the salinity minimum in the Rockall Trough is poorly defined. Long-term monitoring of a section across the northern end of the Trough has shown little change in the properties of LSW since 1975 (D. J. Ellett, manuscript in preparation). The variation in salinity at a depth of 1,600 m has been less than ± 2 standard deviations from the 1975–78 mean. In 1990, however, there was an unprecedented freshening and cooling of up to 7 standard deviations (Ellett, manuscript in preparation). If we relate this to the changes in the Labrador Sea we see (Fig. 2a) that the only time there has been such a long period with little or no change in salinity, followed by a sudden marked freshening, was the period leading up to the renewal in 1972. This suggests that the effects of the renewal of LSW in the early 1970s have just reached the Rockall Trough, and from this we can deduce that there is a circulation time of 18–19 years between the Labrador Sea and the Rockall Trough.

To expand the discussion to cover the changes in water properties over the entire water column, we have made comparisons with previous data sets. The CONVEX-91 north and south sections lay close to the Transient Tracers in the Ocean (TTO) 55° N (1981) and Erika Dan 53.5° N (1962) sections respectively (Fig. 1). We have linearly interpolated each pair of sections onto a common grid with points 40 m apart vertically

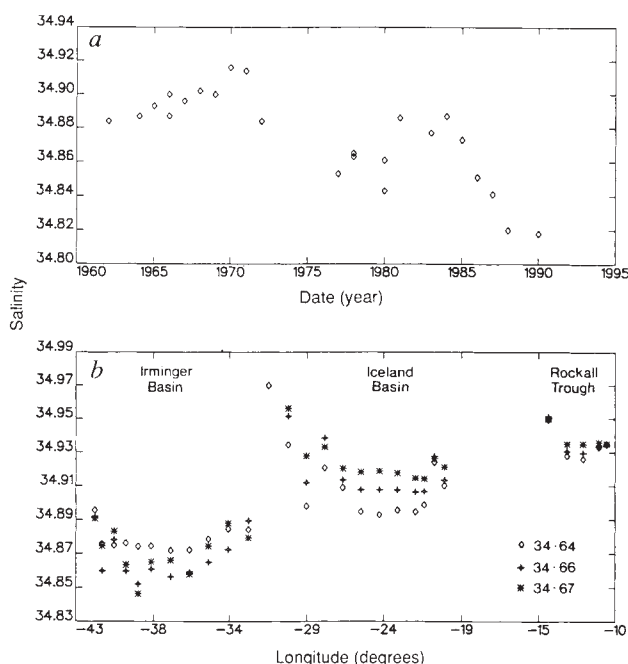


FIG. 2 a, Salinity on the density surface $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$ in the central Labrador Sea from Lazier (personal communication). b, Salinity on the density surfaces $\sigma_{1.5} = 34.64 \text{ kg m}^{-3}$, $\sigma_{1.5} = 34.66 \text{ kg m}^{-3}$, $\sigma_{1.5} = 34.67 \text{ kg m}^{-3}$ from the CONVEX-91 northern section.

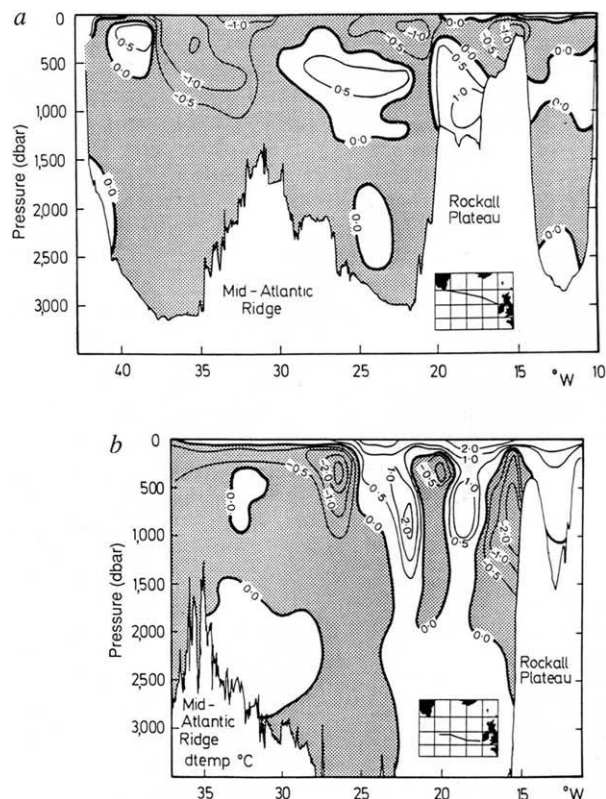


FIG. 3 Smoothed temperature difference (°C) between (a) CONVEX-91 north section and the TTO 55° N section from 1981, and (b) CONVEX-91 south section and the Erika Dan 53.5° N data from 1962. Shaded areas indicate negative temperature difference.

and 0.25° of longitude apart horizontally, and we have subtracted the earlier temperatures from CONVEX-91 temperatures (Fig. 3) (the west sections are not shown). Care must be taken in interpreting Fig. 3b because the CTD stations at the eastern end were not worked in the same location (Fig. 1), and some effect from the different topography might account for the eddy-like structure seen between $15\text{--}28^\circ$ W. Nevertheless the general trend indicated is of a decrease in temperature. Cooling seems to be more pronounced and more extensive in the west than the east; for example, since 1981 (Fig. 3a) there has been a decrease of more than -1°C in the upper layers to the west, compared with warming in the intermediate water to the east of the Mid-Atlantic Ridge. Ignoring measurements above 50 m (the mixed layer depth) the mean temperature difference of -0.15°C between 1991 and 1981 shows that cooling has been much greater than since 1962 (-0.08°C). Previous results showed^{10,11} that the subpolar North Atlantic had warmed between the early 1950s and early 1970s, whereas since then there has been cooling⁶⁻⁸. It would now appear that cooling has continued to temperatures lower than those of 1962. Comparison of the CONVEX-91 west section (Fig. 1) with data from Erika Dan (1962) and Knorr (1983) in a similar fashion (but on a 0.25° latitudinal grid) shows that there has been marked cooling since 1962, averaging -0.42°C , whereas since 1983 there has been no change, presumably because the effects of the first renewal of LSW had already reached the Irminger Basin.

The renewal of LSW during 1972 now seems to have reached the Rockall Trough, with a consequent cooling and freshening across the subpolar North Atlantic. More recently formed LSW, however, is considerably colder and fresher than previously formed LSW and has not yet spread beyond the Irminger Basin.

Given that this was formed in the late 1980s and that our estimated time for previous changes to reach the Rockall Trough was $\sim 18\text{--}19$ years, we predict that the next decade will see continued cooling and freshening of the subpolar gyre as the latest form of LSW circulates around the North Atlantic. Coupled ocean-atmosphere models¹³⁻¹⁵ show surface cooling and freshening of the subpolar North Atlantic in response to greenhouse warming. The freshening dominates in such a way as to suppress convective water mass formation. Our measurements show that cooling is predominant, resulting in renewed convection, and is therefore somewhat at odds with the model results. □

Received 27 April, accepted 25 September 1992.

1. Gould, W. J. *Cruise Report No. 230* (Institute of Oceanographic Sciences Deacon Laboratory, 1992).
2. Clarke, R. A. & Gascard, J. C. *J. phys. Oceanogr.* **13**, 1764-1778 (1983).
3. Talley, L. D. & McCartney, M. S. *J. phys. Oceanogr.* **12**, 1189-1205 (1982).
4. Wallace, D. W. R. & Lazier, J. R. N. *Nature* **332**, 61-63 (1988).
5. Lazier, J. R. N. & Gershey, R. M. *WOCE News* **11**, 5-7 (1991).
6. Brewer, P. G. *et al. Science* **222**, 1237-1239 (1983).
7. Swift, J. H. in *Climate Processes and Climate Sensitivity* (eds Hansen, J. E. & Takahashi, T.) 39-47 (AGU, Washington DC, 1984).
8. Swift, J. H. in *Glaciers, Ice-Sheets, and Sea-Level Effect of a CO_2 -induced Climate Change* (ed. Lingle, C. S.) 129-138 (Department of Energy, Washington DC, 1985).
9. Lazier, J. R. N. *Atmos. Ocean* **18**, 227-238 (1980).
10. Levitus, S. *J. geophys. Res.* **94**, 6091-6131 (1989).
11. Levitus, S. *J. geophys. Res.* **94**, 16125-16131 (1989).
12. Broecker, W. S., Peteet, D. M. & Rind, D. *Nature* **315**, 21-26 (1985).
13. Stouffer, R. J., Manabe, S. & Bryan, K. *Nature* **342**, 660-662 (1989).
14. Manabe, S., Bryan, K. & Stouffer, R. J. *J. phys. Oceanogr.* **20**, 722-749 (1990).
15. Mikolajewicz, U., Santer, B. D. & Maier-Reimer, E. *Nature* **345**, 589-593 (1990).

ACKNOWLEDGEMENTS. The measurements were made on RRS *Charles Darwin* and we thank the Master and crew for their assistance. The quality of the CTD data was assured by Peter Saunders, Mark Brandon, Sheldon Bacon, Pat Gwilliam and Chris Paulson. We also thank John Lazier for additional information on the properties of LSW, and Raymond Pollard for constructive criticism of our manuscript.

A refractory HIMU component in the sources of island-arc magma

Rosemary Hickey-Vargas

Department of Geology, Florida International University, Miami, Florida 33156, USA

THE strontium, neodymium and lead isotopic compositions of most island-arc magmas seem to require a component in the source region with the characteristics of ocean-island basalts (OIB)¹⁻⁵. In contrast, however, the trace-element compositions of arc magmas and OIB do not match: OIB sources are highly enriched in incompatible elements, whereas island-arc basalts are depleted in high-field-strength and light rare-earth elements. Peridotite that has undergone an episode of melting, and hence depletion of incompatible elements, is an appropriate source for island-arc basalts^{6,7}. I have previously proposed that an OIB component with the characteristics of the mantle endmember HIMU⁸ (notably, high U/Pb ratio) is present in arc magmas of the Mariana Island region (Philippine Sea). Here I show, using isotopic comparisons of Philippine Sea arc and basin magmas, including the highly magnesian lavas known as boninites, that this HIMU component is refractory and is selectively incorporated into the depleted mantle lithosphere during extraction of mid-ocean-ridge basalt and back-arc-basin basalt. This process creates the mixture of depleted and OIB-source mantle required for island-arc basalt.

The Philippine Sea region was formed by three successive arc building and back-arc spreading phases, and is therefore ideal for studying the nature and relationship of arc and basin magma sources. Back-arc-basin basalts were erupted 65-40 million years before present (Myr BP) (West Philippine basin), 32-17 Myr BP (Parece Vela and Shikoku basins) and 6-0 Myr BP (Mariana Trough)⁹, and are beginning to be erupted at the Sumisu Rift. Isotope characteristics of these basin magmas are remarkably

uniform (Fig. 1a). On lead isotope diagrams they plot along single linear trends and they share the anomalous Dupal isotope characteristics of mid-ocean-ridge basalts (MORB) from the Indian Ocean, such as elevated $^{207}\text{Pb}/^{204}\text{Pb}$ and $^{208}\text{Pb}/^{204}\text{Pb}$. In contrast, island-arc magmas erupted 45-30 Myr BP (Palau-Kyushu arc), 9-13 Myr BP (West Mariana arc) and 3-0 Myr BP (Mariana arc)⁹ have the Pb-isotope characteristics of the Northern Hemisphere (Fig. 1b), similar to Pacific Ocean mid-ocean-ridge basalt (MORB). Lavas of the Palau-Kyushu arc, in particular, lack evidence of sediment involvement and therefore may reflect the composition of the subarc mantle wedge. Lead isotope ratios in these arc lavas extend from values typical of normal Pacific Ocean MORB (NMORB) towards HIMU⁸, the OIB source component characterized by highest $^{206}\text{Pb}/^{204}\text{Pb}$ (Fig. 1b).

In terms of plate tectonics, the Indian Ocean character of the Philippine Sea basins is not surprising. Adjacent spreading centres in the Japan Sea and South China basin also show Dupal characteristics¹⁰⁻¹², and palaeomagnetic data indicate the Philippine plate originated in the Southern Hemisphere¹³. The alternation of two different mantle signatures in the basin and island arc lavas is, however, difficult to explain within the context of existing arc models, because such models call on either a common asthenospheric source for basin and arc magma¹⁻³, or the derivation of arc magmas from the mantle residue of MORB or back-arc-basin basalt (BABB) generation^{6,7,14}. One possible explanation is that the arc magmas include lead derived from subducted Pacific Ocean crust, including both MORB sea floor and OIB seamounts¹⁵⁻¹⁹, but excluding pelagic sediments¹⁷⁻¹⁹. Detailed consideration of Palau-Kyushu arc magmatism shows, however, that this model is implausible for several reasons. Palau-Kyushu arc lavas with both high and low $^{206}\text{Pb}/^{204}\text{Pb}$ were erupted concurrently at the same sites¹⁷⁻¹⁹, requiring that distinct fluids derived from subducted normal MORB sea floor and OIB seamounts be present at the same location. The higher $^{206}\text{Pb}/^{204}\text{Pb}$ values are found in boninites compared with

tholeiitic basalts¹⁷⁻¹⁹, requiring that boninite magma selectively incorporate seamount-derived fluids. Furthermore, seamounts located on Cretaceous Pacific crust near the Mariana-Bonin Trench vary widely in isotopic character²⁰, whereas magmas from the entire length of the Palau-Kyushu arc form a single trend directed toward the HIMU OIB endmember.

As an alternative, I suggest that the NMORB and HIMU mantle components, together with EM1 and EM2, were original constituents of the asthenosphere tapped by the basin basalts. Like Indian Ocean spreading ridges, the MORB-source asthenosphere may be contaminated with enriched OIB-source components derived from delaminated subcontinental lithosphere, ancient subducted sediments and subducted basaltic crust^{8,21,22} (Fig. 2). During melting to form the basin magmas, the more fusible, enriched endmembers⁸ are selectively extracted, leaving a residue consisting of depleted MORB mantle together with some fraction of the HIMU component. The depleted NMORB-HIMU mixture is cycled into the mantle lithosphere⁷ and becomes the peridotite source for island-arc magmatism (Fig. 2).

Supporting evidence that the HIMU component is refractory is that its signature is most evident in boninites from the Palau-Kyushu arc (Fig. 3). Among arc magma types, boninites have many features indicating an origin by melting of extremely depleted peridotite, such as unusually low incompatible element abundances, high $\text{Al}_2\text{O}_3/\text{TiO}_2$ ratios, high Mg-numbers and Cr-rich spinels²³⁻²⁵. When compared with tholeiitic arc basalts from the same locale, boninites invariably have lower $^{143}\text{Nd}/^{144}\text{Nd}$ and higher $^{206}\text{Pb}/^{204}\text{Pb}$ ^{17-19,25,26}, characteristics of the HIMU component (Fig. 3a). When boninites from throughout the Palau-Kyushu arc are compared with each other, however, specimens with signs of the least depleted source (lowest $\text{Al}_2\text{O}_3/\text{TiO}_2$) have the highest $^{206}\text{Pb}/^{204}\text{Pb}$ (Fig. 3b). My conclusion is that the HIMU component is refractory, but that increasing depletion of peridotite either removes this component or permits other materials, such as fluids from the subducted basaltic crust, to dominate the Pb and Nd isotopic signature.

For island-arc magmas generally, a source consisting of depleted NMORB-source mantle and HIMU explains many of

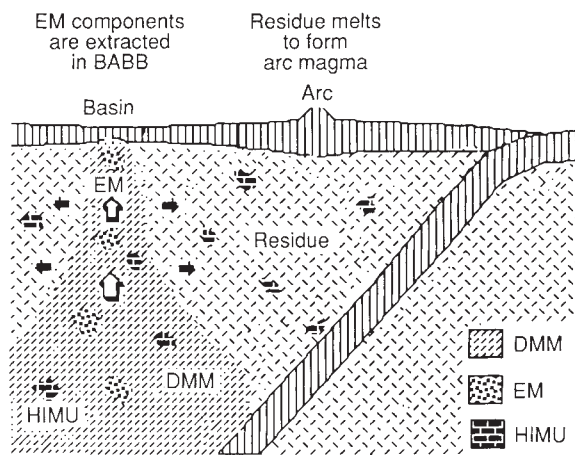


FIG. 2 Possible relationships between four mantle endmembers⁸: EM1, EM2, HIMU and DMM (depleted MORB mantle) during melting to form back arc basin and island arc magma. Fertile asthenosphere includes enriched endmembers EM1 and EM2, and HIMU together with normal MORB source mantle. Melting selectively extracts the enriched (EM) components in the MORB-like basin basalts, leaving a residue of depleted MORB source peridotite and HIMU.

their puzzling isotopic features. The HIMU endmember can produce the lower than MORB-like $^{143}\text{Nd}/^{144}\text{Nd}$ and high or variable $^{206}\text{Pb}/^{204}\text{Pb}$ ratios in arc magmas that lack evidence for sediment involvement (Fig. 4). Further contamination of these sources by sediments will result in lower $^{143}\text{Nd}/^{144}\text{Nd}$ and higher $^{207}\text{Pb}/^{204}\text{Pb}$. A second observation, that many arc magmas have $^{87}\text{Sr}/^{86}\text{Sr}$ ratios too low to include a significant contribution from subducted crust previously altered by contact with sea water^{1,2,27}, may also be explained, as the mixing array of depleted MORB mantle and HIMU has unusually low $^{87}\text{Sr}/^{86}\text{Sr}$.

Although an overall depleted mantle source is consistent with the trace-element features of arc magmas, the trace element signature of an additional, previously melted HIMU component is not entirely clear. HIMU, as sampled by OIB, is enriched in

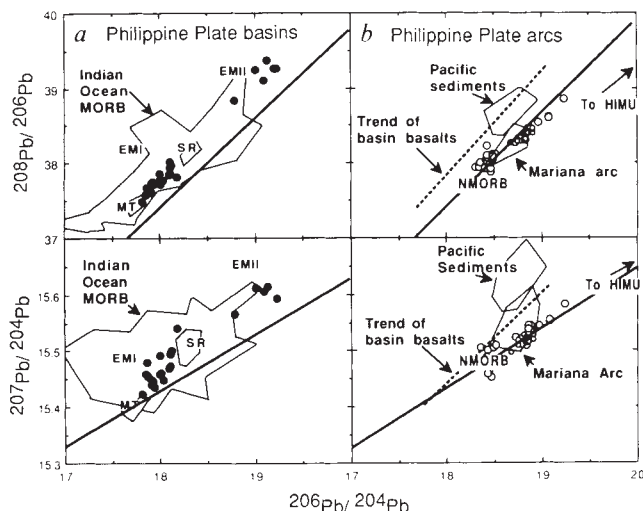


FIG. 1 $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{207}\text{Pb}/^{204}\text{Pb}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$ for Philippine plate basin basalts (a) and arc lavas (b). Filled circles in (a) are data for the West Philippine, Parece Vela and Shikoku basins⁵; MT is the field for N-MORB lavas of the Mariana Trough²⁹, and SR for Sumisu Rift³⁰. Estimated positions of the OIB endmembers of Hart⁸ and the Northern Hemisphere regression line are also shown. Open circles in (b) are data for the Palau-Kyushu arc^{5,17-19,31}, smaller symbols are recent data for ODP sites 782 and 786¹⁹. Fields for the Mariana arc^{15,16,32} and Pacific Plate pelagic sediments^{15,16,32} are also shown.

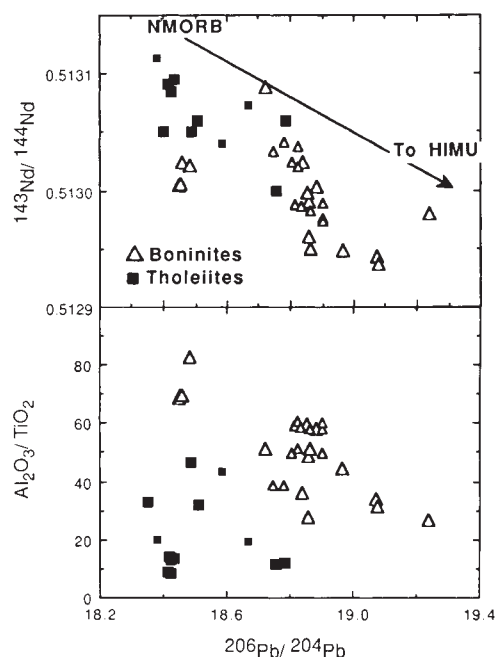


FIG. 3 $^{143}\text{Nd}/^{144}\text{Nd}$ and $\text{Al}_2\text{O}_3/\text{TiO}_2$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$ for boninites^{17-19,31} and tholeiitic rocks^{5,17-19} of the Palau-Kyushu arc. Smaller symbols are recent data for ODP sites 782 and 786¹⁹.

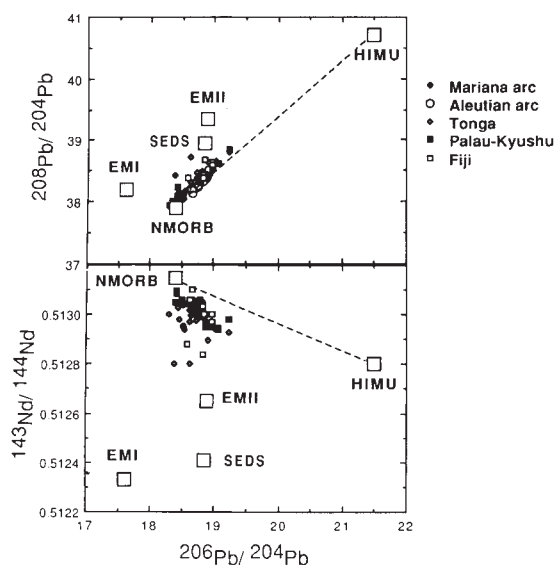


FIG. 4 $^{208}\text{Pb}/^{204}\text{Pb}$ and $^{143}\text{Nd}/^{144}\text{Nd}$ plotted against $^{206}\text{Pb}/^{204}\text{Pb}$ showing relationship of island-arc lavas from the Mariana arc^{15,16,32}, the Aleutian arc^{2,33}, Tonga¹⁴ and the Palau-Kyushu arc^{5,17-19,31} to mixtures of NMORB mantle and the HIMU mantle endmember (dashed line). Approximate values for enriched mantle endmembers EMI and EMII⁸ and Pacific sediment (SEDS)^{15,16,32} are also shown.

elements such as Nb, Zr, La and Nd relative to Rb, Pb, Ba and Sr²⁸, whereas arc magmas show precisely the opposite pattern. The requisite contribution from HIMU must be small compared with that from other source materials, and/or early melting of the HIMU component during basin magmatism must alter these incompatible element relationships. According to Fig. 4, as much as 86% of the Nd, but only 26% of the Pb in arc magmas comes from the HIMU component rather than NMORB peridotite. Because these two components, as sampled by OIB, have demonstrably the same Pb/Nd ratio⁸, their trace element abundances have not been maintained in the sources of island-arc basalts. An additional high-Pb/Nd component, such as fluids expelled from subducted MORB, may complete the picture. □

Received 26 May; accepted 18 September 1992.

1. Stern, R. J. & Ito, E. *J. Volcanol. geotherm. Res.* **18**, 461-482 (1983).
2. Morris, J.D. & Hart, S.R. *Geochim. cosmochim. Acta* **47**, 2015-2030 (1983).
3. Gill, J.B. *Earth planet. Sci. Lett.* **68**, 443-458 (1984).
4. Lin, P. N., Stern, R. J., Morris, J. & Bloomer, S. H. *Contr. Miner. Petrol.* **105**, 381-392 (1990).
5. Hickey-Vargas, R. *Earth planet. Sci. Lett.* **107**, 290-304 (1991).
6. Kay, R. W. *J. Geol.* **88**, 497-522 (1980).
7. McCulloch, M. T. & Gamble, J. A. *Earth planet. Sci. Lett.* **102**, 358-374 (1991).
8. Hart, S. R. *Earth planet. Sci. Lett.* **90**, 273-296 (1988).
9. Hussong, D. M. & Uyeda, S. *Init. Rep. DSDP* **60**, 909-929 (1981).
10. Tu, K., Flower, M., Carlson, R., Zhang, M. & Xie, G. *Geology* **19**, 567-569 (1991).
11. Tatsumoto, M. & Nakamura, Y. *Geochim. cosmochim. Acta* **55**, 3697-3708 (1991).
12. Chung, S.-L. & Sun, S.-S. *Earth planet. Sci. Lett.* **109**, 133-145 (1992).
13. Seno, T. & Maruyama, S. *Tectonophysics* **102**, 53-84 (1984).
14. Ewart, A. & Hawkesworth, C. J. *J. Petrol.* **28**, 495-530 (1987).
15. Woodhead, J. D. & Fraser, D. G. *Geochim. cosmochim. Acta* **49**, 1925-1930 (1985).
16. Woodhead, J. D. *Chem. Geol.* **76**, 1-24 (1989).
17. Hickey-Vargas, R. & Reagan, M. K. *Contr. Miner. Petrol.* **97**, 497-508 (1987).
18. Hickey-Vargas, R. in *Boninites* (ed. Crawford, A. J.) 339-356 (Unwin-Hyman, London, 1989).
19. Pearce, J. A. et al. *Proc. ODP Sci. Res.* **125**, 237-261 (1992).
20. Staudigel, H. et al. *Earth planet. Sci. Lett.* **102**, 14-44 (1991).
21. Ito, E., White, W. M. & Gopel, C. *Chem. Geol.* **62**, 157-176 (1987).
22. Castillo, P. *Nature* **336**, 667-670 (1988).
23. Sun, S.-S. & Nesbitt, R. W. *Geology* **6**, 689-693 (1978).
24. Cameron, W. E., Nisbet, E. G. & Dietrich, V. J. *Nature* **280**, 550-552 (1979).
25. Hickey, R. L. & Frey, F. A. *Geochim. cosmochim. Acta* **46**, 2099-2115 (1982).
26. Cameron, W. E., McCulloch, M. T. & Walker, D. A. *Earth planet. Sci. Lett.* **65**, 75-89 (1983).
27. DePaolo, D. J. & Johnson, R. W. *Contr. Miner. Petrol.* **70**, 367-379 (1979).
28. Weaver, B. L. *Earth planet. Sci. Lett.* **104**, 381-397 (1991).
29. Volpe, A. M., Macdougall, J. D., Lugmair, G. W., Hawkins, J. W. & Lonsdale, P. *Earth planet. Sci. Lett.* **100**, 251-264 (1990).
30. Hochstaedter, A. G., Gill, J. B. & Morris, J. D. *Earth planet. Sci. Lett.* **100**, 195-209 (1990).
31. Stern, R. J., Morris, J., Bloomer, S. H. & Hawkins, J. W., *Geochim. cosmochim. Acta* **55**, 1467-1481 (1991).
32. Meijer, A. *Geol. Soc. Am. Bull.* **87**, 1358-1369 (1976).
33. Kay, R. W., Sun, S.-S. & Lee-Hu, C.-N. *Geochim. cosmochim. Acta* **42**, 263-273 (1978).

Spring phytoplankton blooms in the absence of vertical water column stratification

David W. Townsend*, Maureen D. Keller*, Michael E. Sieracki* & Steven G. Ackleson†

* Bigelow Laboratory for Ocean Sciences, McKown Point, W. Boothbay Harbor, Maine 04575, USA

† Lockheed Engineering and Science Co., 2400 Nasa Road One, Mail Code C102, Houston, Texas 77258, USA

THE spring phytoplankton bloom in temperate and boreal waters represents a pulsed source of organic carbon that is important to ecosystem productivity¹ and carbon flux² in the world ocean. It is widely accepted that the seasonal development of a thermocline, in combination with increasing solar elevation in spring, is requisite for the development of the bloom in shelf and open ocean environments³⁻⁷. Here we report results for the offshore waters of the Gulf of Maine which suggest that the spring bloom can precede the onset of vertical water column stability, and may even be a contributing factor in the development of the thermocline. Deep penetration of light in relatively clear, late-winter waters, and weak, or absent, wind-driven vertical mixing, appear to support cell growth rates that overcome the vertical excursion rates in the neutrally stable water column, leading to a bloom. Phytoplankton forms typical of a spring bloom, including gelatinous colonies and chains, may contribute to the cells' ability to maintain a vertical position in a water column lacking stability.

During the winter months in mid- to high-latitude waters, low solar elevations and *in situ* marine light levels, along with deep turbulent vertical mixing of the upper water column, maintain phytoplankton production at its lowest levels of the year. This is despite high concentrations of inorganic nutrients in the surface waters that result from winter convective mixing with deep waters. Many have argued that as the seasonal thermocline develops and phytoplankton cells in the upper mixed layer become isolated above some 'critical depth', phytoplankton photosynthesis exceeds respiration and the bloom commences³⁻⁷. Field studies have shown that a depth-averaged, vertically integrated irradiance value of about 20.9 W m^{-2} defines this critical depth^{4,8-14}. The depth of this critical light intensity increases with solar elevation as spring progresses and, at some point, it intersects with the bottom in inshore waters, or with the depth of the developing upper mixed layer created above the thermocline in deeper offshore waters on the shelf and in the open ocean³⁻¹⁴.

The amount of light available for photosynthesis that reaches a particular depth (E_z) is described by

$$E_z = E_0(1-r) \exp \left[\int_0^z -K_z dz \right] \quad (1)$$

where E_0 is the solar radiation reaching the sea surface (W m^{-2}), r is the reflectance at the air-sea boundary ($r \approx 0.021$ averaged over the visible spectrum), K_z is the depth-dependent diffuse attenuation coefficient (m^{-1}), and z is the depth (m). As E_0 is the integrated irradiance over the entire visible light spectrum, K_z is expected to change with depth as the ocean water mixture (the combination of pure water and any absorbing particulate and dissolved material) selectively removes some wavelengths near the surface and transmits other wavelengths to greater depths. The depth-dependence of K_z has been described in terms of an arc-tangent model¹⁵

$$E_z = E_0(1-r) \exp(-k_1 z) [1 - k_2 \tan^{-1}(k_3 z)] \quad (2)$$

where the coefficients k_1 , k_2 and k_3 are derived from measured irradiance profiles. The daily irradiance at the sea surface in the

absence of cloud cover may be expressed as

$$E_0 = \alpha + \beta \sin \omega t \quad (3)$$

where α is the daily irradiance during the winter solstice; β is the difference in daily irradiance between the winter and summer solstice; $\omega = \pi/365 \text{ d}^{-1}$; and t is time (d) past the winter solstice (21 December). For the Gulf of Maine, $\alpha \approx 59.9 \text{ W m}^{-2}$ and $\beta \approx 169.4 \text{ W m}^{-2}$ (ref. 16). The depth-averaged, vertically integrated irradiance in the sea, E^* , between the ocean surface and the bottom of a surface mixed layer at depth z_m describes the average amount of light experienced by a phytoplankton population within the mixed layer, and is defined as

$$E^* = 1/z_m \int_0^{z_m} E_z dz \quad (4)$$

Vertical light profiles were measured in the offshore Gulf of Maine 21–30 April 1991 using a LiCor submersible spherical radiometer. These data were assumed to represent pre-bloom, clear water conditions in early spring; phytoplankton chlorophyll *a* values at this time were less than 1.0 mg m^{-3} . The corresponding k coefficients in equation (2) were derived from irradiance profiles measured at these non-bloom Gulf of Maine stations; $k_1 = 0.089 \text{ m}^{-1}$, $k_2 = 0.381$, $k_3 = 1.72 \text{ m}^{-1}$. E^* was then computed at increments of 1 m between the surface and 100 m for each day between the winter solstice (21 December) and the summer solstice (21 June). On each day the critical depth was identified as that for which $E^* = 20.9 \text{ W m}^{-2}$. These simulations were then compared with field observations of water column

structure during the spring phytoplankton bloom we observed during research cruises in 1990 and 1992.

A spring bloom was observed in offshore Massachusetts Bay in the western Gulf of Maine in March of 1990, which was occurring in the absence of vertical stratification (Fig. 1*a, b*). The water column was vertically isopycnal from the surface to 70 m, with no change in density (σ_t) except very near the bottom. The chlorophyll *a* concentrations were $\sim 4 \text{ mg m}^{-3}$, and the diffuse attenuation coefficients (K_z) ranged from $\sim 0.4 \text{ m}^{-1}$ near the surface to 0.25 m^{-1} at 40 m. The light field in early March, based on our calculations for non-blooming Gulf of Maine waters in April of 1991, would dictate a pre-bloom critical depth of about 38 m (where $E^* = 20.9 \text{ W m}^{-2}$), but the bloom was progressing despite the lack of any vertical water column stratification shallower than 70 m. The deep chlorophyll fluorescence in Fig. 1*a, b* seems to be the result of sinking aggregates of phytoplankton cells, as described earlier^{17–21}.

In early April of 1992 the spring bloom was underway throughout most of the Gulf of Maine, and was occurring at stations with little or no vertical stratification. Density differences in the upper water column ($\Delta\sigma_t$) were 0.04 kg m^{-3} from 0 to 100 m at the stations in Fig. 2*a* and *b*, whereas the calculated critical depth in these waters for this date was about 43 m. The chlorophyll *a* concentrations exceeded 2.5 mg m^{-3} . Not all stations were neutrally stable, and the bloom was also underway at stations with salinity stratification at depths near, or less than, the critical depth, as shown in Fig. 2*c*. The deep extension of the chlorophyll fluorescence in Fig. 2*a* and *b* also shows evidence of sinking aggregates of cells.

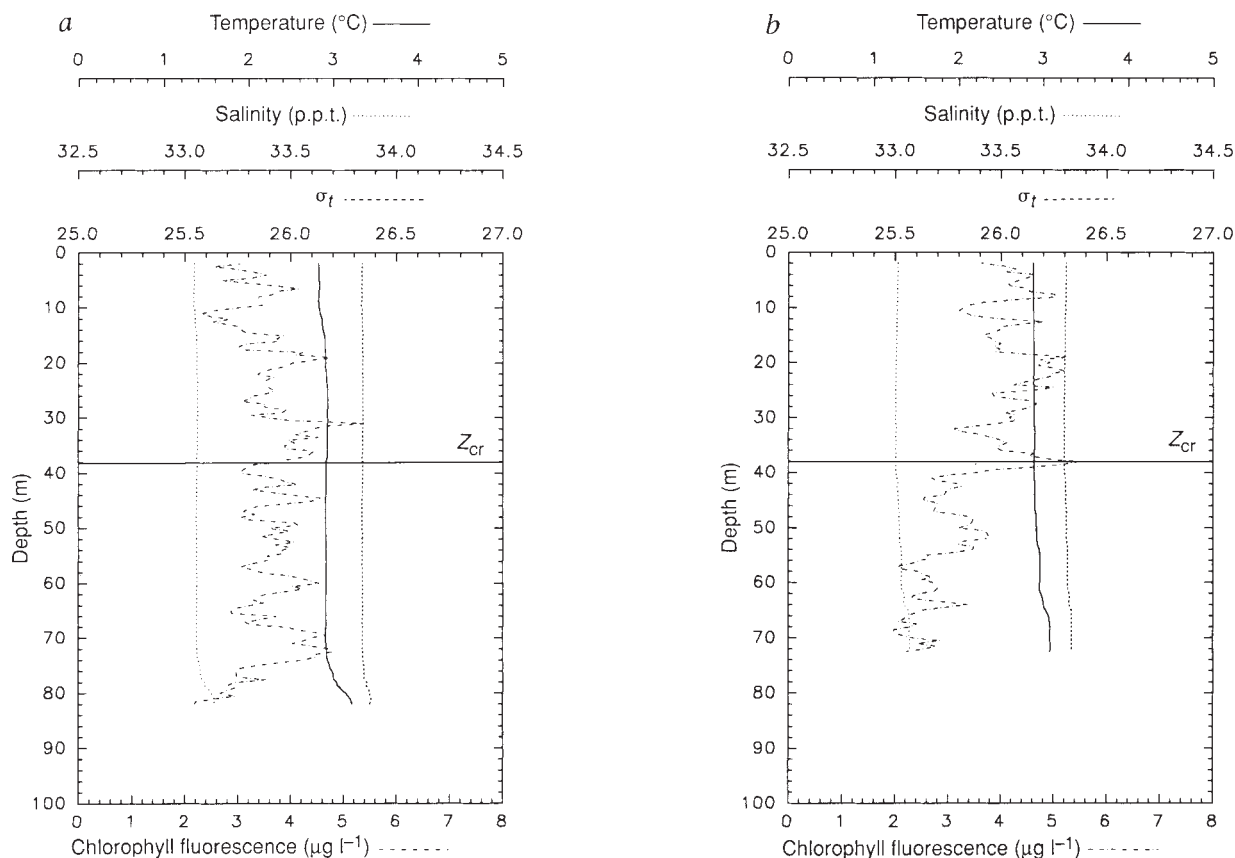


FIG. 1 Vertical profiles showing bloom levels of chlorophyll in the absence of vertical stratification and a critical depth (Z_{cr}) of 38 m. Temperature ($^{\circ}\text{C}$), salinity (p.p.t.), density (σ_t , kg m^{-3}) and *in situ* chlorophyll fluorescence ($\text{mg chlorophyll } a \text{ m}^{-3}$) were measured at stations in Massachusetts Bay in the western Gulf of Maine on 6 March 1990 using a Neil Brown Mark III CTD with a Sea Tech *in situ* fluorometer. The error in the *in situ* chlorophyll fluorescence was about $\pm 1 \text{ mg m}^{-3}$ chlorophyll *a*, based on acetone extractions of water samples. *a*, Measurements taken at $42^{\circ} 19.9' \text{N}$, $70^{\circ} 25.8' \text{W}$,

about 20 nautical miles west of Boston, bottom depth 86 m, and *b*, at $42^{\circ} 18.4' \text{N}$, $70^{\circ} 32.9' \text{W}$, about 18 nautical miles west of Boston, bottom depth 74 m. The dominant phytoplankton species at these two stations, determined from microscopic examinations of water samples, were *Detonula confervacea*, *Thalassiosira pseudonana* and *Chaetoceros socialis*²⁷. The spiked fluorescence traces are apparently the result of flocculated cells, which can occur at densities of ~ 10 per litre (ref. 21). The horizontal lines at 38 m (Z_{cr}) indicate the critical depth.

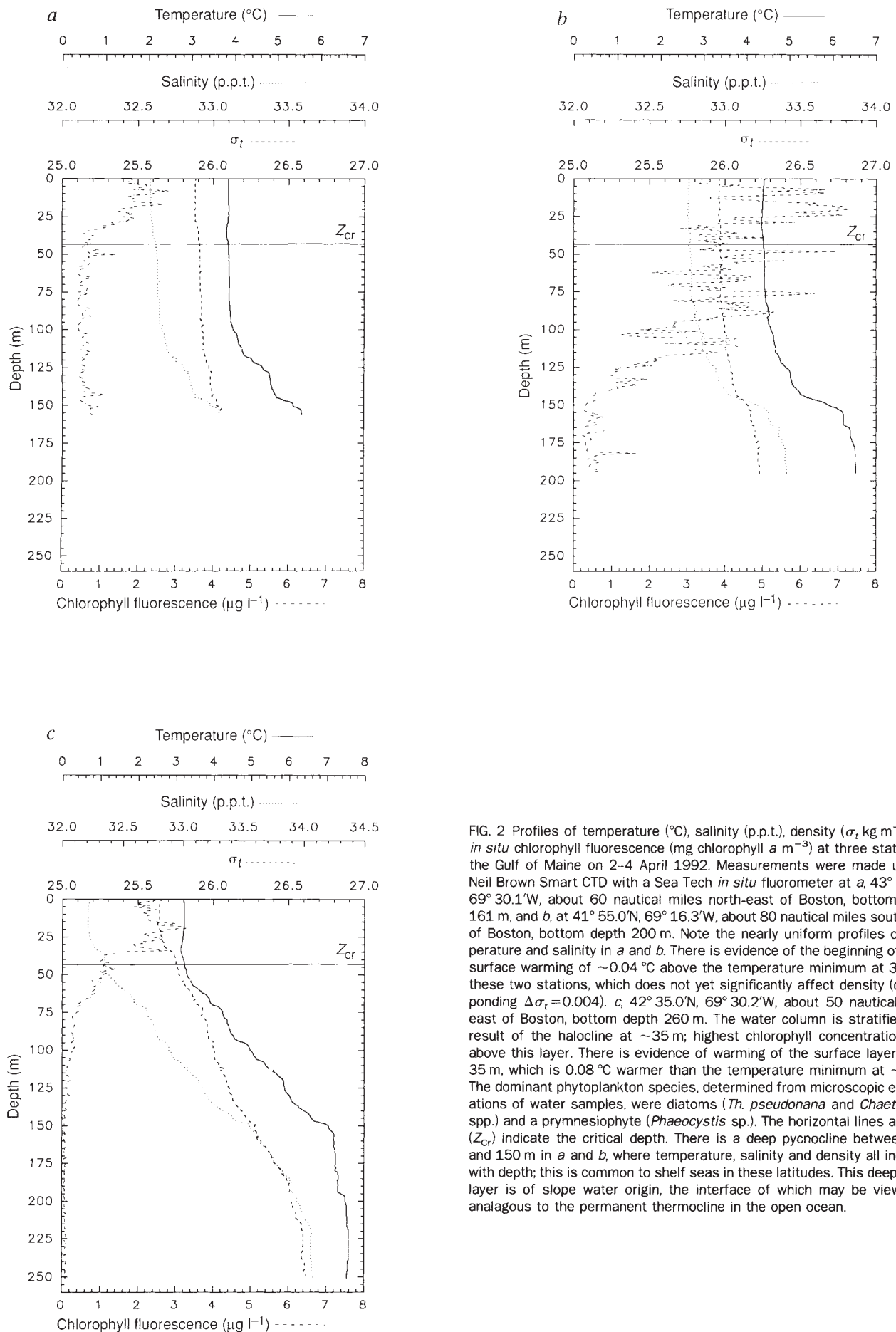


FIG. 2 Profiles of temperature ($^{\circ}\text{C}$), salinity (p.p.t.), density (σ_t kg m^{-3}) and *in situ* chlorophyll fluorescence ($\text{mg chlorophyll } a \text{ m}^{-3}$) at three stations in the Gulf of Maine on 2–4 April 1992. Measurements were made using a Neil Brown Smart CTD with a Sea Tech *in situ* fluorometer at *a*, 43° 20.0'N, 69° 30.1'W, about 60 nautical miles north-east of Boston, bottom depth 161 m, and *b*, at 41° 55.0'N, 69° 16.3'W, about 80 nautical miles south-east of Boston, bottom depth 200 m. Note the nearly uniform profiles of temperature and salinity in *a* and *b*. There is evidence of the beginning of slight surface warming of $\sim 0.04^{\circ}\text{C}$ above the temperature minimum at 30 m at these two stations, which does not yet significantly affect density (corresponding $\Delta\sigma_t = 0.004$). *c*, 42° 35.0'N, 69° 30.2'W, about 50 nautical miles east of Boston, bottom depth 260 m. The water column is stratified as a result of the halocline at ~ 35 m; highest chlorophyll concentrations are above this layer. There is evidence of warming of the surface layer above 35 m, which is 0.08°C warmer than the temperature minimum at ~ 35 m. The dominant phytoplankton species, determined from microscopic examinations of water samples, were diatoms (*Th. pseudonana* and *Chaetoceros* spp.) and a prymnesiophyte (*Phaeocystis* sp.). The horizontal lines at 43 m (Z_{cr}) indicate the critical depth. There is a deep pycnocline between 125 and 150 m in *a* and *b*, where temperature, salinity and density all increase with depth; this is common to shelf seas in these latitudes. This deep-water layer is of slope water origin, the interface of which may be viewed as analogous to the permanent thermocline in the open ocean.

Although our results are from a continental shelf sea, it is important to note that upper water column physical processes over deep shelf waters are the same as in the open ocean, where there are anecdotal accounts of similar phytoplankton blooms in the absence of vertical stratification. Earlier workers in the North Atlantic have observed spring phytoplankton blooms that apparently began before the development of the thermocline, prompting the suggestion of transient thermoclines^{22–24}. One study reported that a top-to-bottom temperature difference of as little as 0.2 °C in shelf waters apparently provided sufficient stabilization of the water column to allow the bloom to commence¹¹. In the Gulf of Maine, a density difference of 0.1–0.2 $\Delta\sigma_t$ in the top 30 m has been assumed to be required for the initiation of the spring bloom⁵. None of these earlier studies provide conclusive evidence that the development of the thermocline or

water column stability preceded the onset of the bloom, though examples do exist. Our results suggest that vertical water column stability may not be a prerequisite to maintain phytoplankton cells above the critical depth. The deepening penetration of light in the clear winter waters during the spring months, in conjunction with absent, or weak, vertical wind mixing, could maintain cell growth rates that overcome the vertical excursion rates in the neutrally stable water column, thus leading to a bloom. Once begun, the bloom may enhance the warming of the surface waters (Fig. 2) by virtue of the light scattering and absorption properties of the phytoplankton particles^{25,26}. Thus, in some instances, the development of the thermocline may be initiated by, rather than serve as a prerequisite for, the spring phytoplankton bloom. □

Received 28 May; accepted 17 September 1992.

1. Smetacek, V. *Ophelia* **1** (suppl.) 65–76 (1980).
2. JGOFS (Joint Global Ocean Flux Study) North Atlantic Bloom Experiment, International Scientific Symposium, Washington, JGOFS Report No. 7, November 1990.
3. Sverdrup, H. U. *J. Conseil exp. Mer.* **18**, 287–295 (1953).
4. Riley, G. A. *Limnol. Oceanogr.* **2**, 252–270 (1957).
5. Bigelow, H. B., Lillick, L. C. & Sears, M. *Trans. Am. Phil. Soc.* **21**, 149–191 (1940).
6. Atkins, W. R. G. *J. mar. Biol. Assoc. U.K.* **15**, 191–205 (1928).
7. Gran, H. H. & Braarud, T. *J. biol. Bd. Can.* **1**, 279–467 (1935).
8. Barlow, J. P. *J. mar. Res.* **17**, 53–67 (1958).
9. Riley, G. A. in *Estuaries* (ed. Lauff, G. H.) 316–326 (Am. Assoc. Adv. Sci. Publ. **83**, Washington DC, 1967).
10. Gieskes, W. W. C. & Kraay, G. W. *Neth. J. Sea Res.* **9**, 166–196 (1975).
11. Pingree, R. D., Holligan, P. M., Mardell, G. T. & Head, R. N. *J. mar. Biol. Assoc. U.K.* **56**, 845–873 (1976).
12. Hitchcock, G. L. & Smayda, T. J. *Limnol. Oceanogr.* **22**, 126–131 (1977).
13. Horn, H. & Paul, L. *Int. Rev. ges. Hydrobiol.* **69**, 507–519 (1984).
14. Townsend, D. W. & Spinrad, R. W. *Cont. Shelf Res.* **6**, 515–529 (1986).

15. Zaneveld, J. R. V. & Spinrad, R. W. *J. geophys. Res.* **85**(C9), 4919–4922 (1980).
16. Sverdrup, H. U., Johnson, M. W. & Fleming, R. H. *The Oceans, Their Physics, Chemistry and General Biology* (Prentice-Hall, Englewood Cliffs, NJ, 1942).
17. Jackson, G. A. *Deep Sea Res.* **37**, 1197–1211 (1990).
18. Wassman, P., Vernet, M., Mitchell, B. G. & Rey, F. *Mar. Ecol. Prog. Ser.* **66**, 183–195 (1990).
19. Riebesell, U. *Mar. Ecol. Prog. Ser.* **69**, 281–291 (1991).
20. Kranck, K. & Milligan, T. G. *Mar. Ecol. Prog. Ser.* **44**, 183–189 (1988).
21. Alldredge, A. L. & Gotschalk, C. C. *Deep Sea Res.* **36**, 159–171 (1989).
22. Colebrook, J. M. *Mar. Biol.* **51**, 23–32 (1979).
23. Colebrook, J. M. *J. Plankton Res.* **4**, 435–462 (1982).
24. Williams, R. & Robinson, G. A. *Ann. Biol. Copenh.* **28**, 57–59 (1973).
25. Lewis, M. R., Cullen, J. J. & Platt, T. *J. geophys. Res.* **88**, 2565–2570 (1983).
26. Mazumder, A., Taylor, W. D., McQueen, D. J. & Lean, D. R. S. *Science* **247**, 312–315 (1989).
27. Townsend, D. W. et al. *Bigelow Laboratory for Ocean Sciences, Tech. Rep.* **76**, 256p (1990).

ACKNOWLEDGEMENTS. We thank S. Corwin, J. Brown, E. Haugen, W. Bellows, T. Cucci, T. Boynton and D. Phinney for assistance. The research was supported by the NSF.

Electrical signalling and systemic proteinase inhibitor induction in the wounded plant

D. C. Wildon*, J. F. Thain*, P. E. H. Minchin†, I. R. Gubb*, A. J. Reilly*, Y. D. Skipper*, H. M. Doherty‡, P. J. O'Donnell‡ & D. J. Bowles‡

* School of Biological Sciences, University of East Anglia, Norwich, NR4 7TJ, UK

† The Horticulture and Food Research Institute of New Zealand Ltd, Lower Hutt, New Zealand

‡ Centre for Plant Biochemistry and Biotechnology, University of Leeds, Leeds LS2 9JT, UK

THE wound response of several plant species involves the activation of proteinase inhibitor (*pin*) genes and the accumulation of *pin* proteins at the local site of injury and systemically throughout the unwounded aerial regions of the plant^{1,2}. It has been suggested that a mobile chemical signal is the causal agent linking the local wound stimulus to the distant systemic response, and candidates such as oligosaccharides³, abscisic acid⁴ and a polypeptide^{5,6} have been put forward. But the speed of transmission is high for the transport of a chemical signal in the phloem. The wound response of tomato plants can be inhibited by salicylic acid⁷ and agents like fusaric acid that affect ion transport⁸, and wounding by heat⁹ or physical injury produces electrical activity that has similarities to the epithelial conduction system¹⁰ used to transmit a stimulus in the defence responses of some lower animals¹¹. Here we design experiments to distinguish between a phloem-transmissible chemical signal and a physically propagated signal based on electrical activity. We show that translocation in the phloem of tomato seedlings can be completely inhibited without effect on the systemic accumulation of *pin* transcripts and *pin* activity, and without hindrance to propagated electrical signals.

Mechanical damage applied to the lamina of one of the cotyledons of a tomato seedling with one expanded leaf (Fig. 1a) leads to electrical activity (Fig. 1b) which can be detected by extracellular electrodes placed on the stem and on the petiole of leaf 1, morphologically the lowest (first-formed) leaf (Fig. 1a). The electrical activity propagates beyond the cotyledon and through the plant at a speed of between 1 and 4 mm s⁻¹. We refer to this activity as a systemic electrical signal. Chilling the petiole of the wounded cotyledon to 3 °C has little effect on the electrical signal recorded on the petiole of leaf 1 (Fig. 1c), a result that is consistent with observations on action potential conduction at low temperature in cold-blooded animals¹² and plants¹³. Chilling a petiole is, however, a recognized means of inhibiting phloem transport¹⁴. In these experiments chilling is necessary, given that the maximum reported speeds of phloem transport¹⁵ (35–250 mm min⁻¹) are such that a chemical signal might exit the wounded cotyledon in a time comparable to that for an electrical signal. As the effects of chilling and their duration are dependent on plant species¹⁴, we analysed the effect of the treatment on tomato seedlings. As shown in Fig. 2, chilling of the cotyledonary petiole to 4 °C completely stops translocation of ¹⁴C-labelled photosynthate through the chilled region for a period of at least ten minutes before a spontaneous recovery. Chilling the cotyledonary petiole to 0.5 °C extends the time for spontaneous recovery to about 25 minutes. Similar results are obtained on chilling the petiole of leaf 1 to 3 °C, except that translocation is inhibited for at least 20 minutes. Given the spontaneous recovery of the petiole tissue from chilling, it was necessary to excise the wounded cotyledon within the period in which phloem transport is inhibited in order to prevent any delayed movement of a chemical signal from the wound-site through the phloem. Leaf excision alone has little effect on *pin* induction in tomato plants⁷.

In the experiments shown in Table 1, *pin* activity was assayed by the inhibitory effect of leaf extracts on chymotrypsin activity (Table 1 legend). Untreated tomato seedlings (line 1) contain negligible *pin* activity and the chymotrypsin is not inhibited.

Chilling of the cotyledonary petiole (line 2) and excision of the cotyledon after chilling (line 3) do not induce detectable systemic *pin* activity. Unwounded seedlings subjected to these conditions (lines 1–3) also did not exhibit any systemic electrical signal, either spontaneously or on excision of the cotyledon. A mechanical wound to the cotyledon (lines 4–6) led to a systemic accumulation of *pin* activity and levels of this activity were identical whether the cotyledon was excised five minutes after wounding or left attached to the plant for 48 hours. The systemic induction of *pin* activity by wounding was correlated on 46 out of 49 occasions with a systemic electrical signal. On one occasion (line 6) the mechanical wound led neither to *pin* activity nor to systemic electrical activity. The induction of systemic *pin* activity by mechanical wounding of the cotyledon was highly reproducible, in contrast to the systemic response of *pin* activity to mechanical wounding of leaf 1 of larger tomato seedlings with two expanded leaves. For these larger seedlings, we observed much greater reproducibility when a heat stimulus was applied to leaf 1. As chilling of the petiole of leaf 1 inhibited phloem translocation for about 20 min, we assayed the effects of chilling the petiole on heat stimulus-induced systemic accumulation of *pin* activity in leaf 2 and on the systemic electrical signal. The data in Table 1 (lines 7–9) show a close correspondence between the systemic *pin* activity and the systemic electrical response;

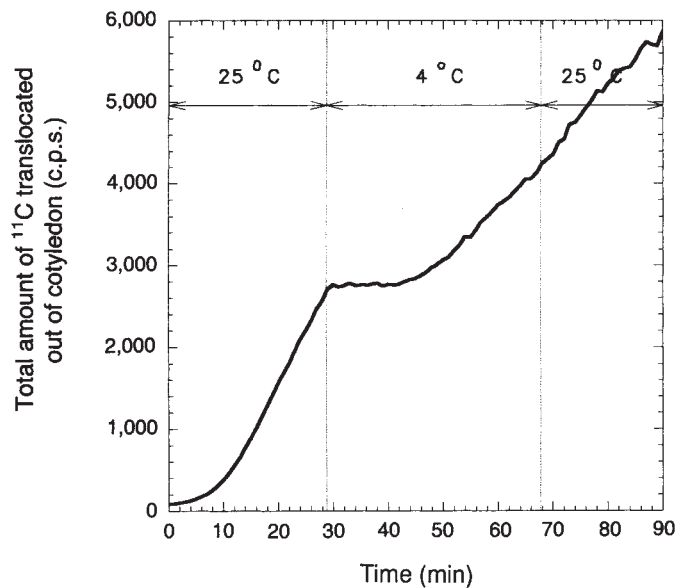


FIG. 2 Translocation of ^{11}C -labelled photosynthate out of one of the cotyledons of a tomato seedling which had been pulse-labelled with $^{11}\text{CO}_2$ at time zero. Radioactivity (corrected for decay) was measured as counts per second in the plant beyond the cooling block, which was located on the petiole of the cotyledon. A 5-mm-long cold block was applied at 29 min and removed at 68 min. Methods as in ref. 14; plant material obtained as in ref. 9.

chilling of the petiole did not affect these responses. Thus, whether the wound stimulus arises from mechanical wounding to the cotyledon, or heat injury to leaf 1, a systemic accumulation of *pin* activity can be correlated with a systemic electrical signal and neither response is affected by the inhibition of phloem translocation.

To confirm that the changes in *pin* activity were due to changes in the level of a characterized gene product of tomato, we followed changes in steady-state levels of *pin 2* messenger RNA by northern analysis. The results are shown in Fig. 3: *pin* transcripts are below the level of detection in leaf 1 of untreated tomato seedlings and also after chilling of the cotyledonary petiole. Even in the presence of a cold block, *pin 2* mRNA levels increase in the unwounded leaf 1 in response to a mechanical wound to the cotyledon, as anticipated, and the increase is substantially greater than the trace increase in level induced by excision of the unwounded cotyledon. Steady-state levels of *pin 2* mRNA were identical whether the cotyledon was excised 5 min after wounding (Fig. 3D) or left attached to the plant for 4 hours (Fig. 3E) and are comparable to those observed in other studies of wound-inducible *pin* mRNA levels^{16,17}. Each mechanically wounded plant also exhibited a systemic electrical signal.

Our data show that the systemic responses of wound-induced *pin* activity and electrical signals occur when phloem translocation is inhibited. Equivalent and high levels of *pin* activity and *pin* transcripts were found in the systemically responding leaf 1, whether the wounded cotyledon was excised rapidly or left attached to the plant. This indicates that the causal agent of the systemic response has exited the wound site within 5 minutes and contrasts with studies on older plants in which systemic *pin* levels increased as the time from wounding to excision increased until 120 minutes after wounding¹⁸. Our results show that the causal agent exits the wounded cotyledon when phloem translocation is blocked, so the systemic effect must either arise from a chemical signal travelling through a route that does not involve the transport of photosynthate, or a physical signal. Movement of a chemical signal could be through the xylem, although this would require a reversal in the direction of transpirational water flow. To analyse this possibility we measured stomatal resistance

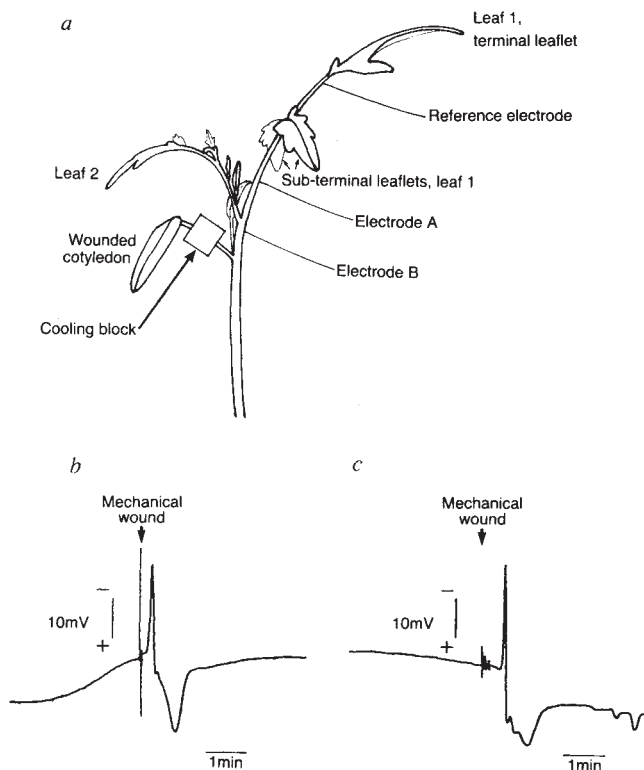


FIG. 1 A typical tomato seedling as used in the mechanical wounding experiments. *a*, Experimental arrangement. The seedling has one expanded leaf, labelled leaf 1, which is morphologically the lowest (first-formed) leaf. *b*, Recording on the petiole of leaf 1; mechanical wound to one of the cotyledons; petiole of cotyledon not chilled. *c*, Recording on the petiole of leaf 1; mechanical wound to one of the cotyledons; petiole of cotyledon chilled (3 °C) from 5 min before wounding until 5 min after wounding.

METHODS. Electrical recordings were made between two thread electrodes (electrode A and the reference electrode), both attached to leaf 1 of a tomato seedling. Recordings (not shown) were also made between the same reference electrode and the thread electrode (B) connected to the stem between the cotyledonary node and leaf 1. A 10-mm cooling block was located on the petiole of the cotyledon and operated to give a petiole temperature of either $\sim 22\text{ }^{\circ}\text{C}$ (*b*) or $\sim 3\text{ }^{\circ}\text{C}$ (*c*). For details of plant material, growth conditions and electrical recording methods, see ref. 9. Details of wounding are given in the legend to Table 1.

TABLE 1 Inhibition of added chymotrypsin by leaf extracts

Treatment	Chymotrypsin activity*		Systemic electrical signal on wounding	Treatment	Chymotrypsin activity*		Systemic electrical signal on wounding
	Mean $\pm$ s.e.	(n)			Mean $\pm$ s.e.	(n)	
1 Untreated plants (data are for leaf 1)	97 $\pm$ 2	(20)	No systemic electrical signal	7 No heat stimulus; petiole of leaf 1 chilled for 15 min before excision of leaf 1 (data are for leaf 2)	102 $\pm$ 4	(12)	No systemic electrical signal
2 No mechanical wound; petiole of cotyledon chilled (3 °C) for 10 min; no excision (data are for leaf 1)	95 $\pm$ 1	(10)	No systemic electrical signal	8 Heat stimulus to tip of leaf 1; petiole of leaf 1 not chilled; leaf 1 excised 10 min after heat stimulus applied (data are for leaf 2)	20 $\pm$ 2	(12)	Systemic electrical signal observed in every replicate
3 No mechanical wound; petiole of cotyledon chilled (3 °C) for 10 min before excision of the cotyledon (data are for leaf 1)	94 $\pm$ 3	(10)	No systemic electrical signal	9 Heat stimulus to tip of leaf 1; petiole of leaf 1 chilled (3 °C) from 5 min before heat stimulus until excision of leaf 1 at 10 min after heat stimulus (data are for leaf 2)	25 $\pm$ 7	(12)	Systemic electrical signal observed in every replicate (for one replicate no <i>pin</i> activity was detected)
4 Mechanical wound to one of the cotyledons; petiole of cotyledon not chilled; no excision (data are for leaf 1)	21 $\pm$ 4	(10)	Systemic electrical signal observed in every replicate	10 Mechanical wound to one of the cotyledons; petiole of cotyledon not chilled; cotyledon excised adjacent to the stem when an electrical signal was just starting to be detected at an electrode located on the cotyledonary petiole near the lamina; mean time from wound to excision 28 s (data are for leaf 1)	91 $\pm$ 4	(8)	No systemic electrical signal
5 Mechanical wound to one of the cotyledons; petiole of cotyledon not chilled; wounded cotyledon excised 5 min after mechanical wounding (data are for leaf 1)	28 $\pm$ 6	(20)	Systemic electrical signal observed in every replicate (for three replicates little or no <i>pin</i> activity was detected)				
6 Mechanical wound to one of the cotyledons; petiole of cotyledon chilled (3 °C) from 5 min before mechanical wounding until excision of wounded cotyledon at 5 min after mechanical wounding (data are for leaf 1)	27 $\pm$ 6	(20)	Systemic electrical signal observed in all except one replicate; for this replicate no <i>pin</i> activity was detected				

A mechanical wound was produced by crushing an area of $\sim 1 \text{ cm}^2$ at the distal end of the cotyledon of a tomato seedling at the stage of development shown in Fig. 1a, followed immediately by crushing a similar area on the proximal side of the first wound. The heat stimulus⁹ was provided by a stainless steel block (contact area, 3 cm^2) heated in a boiling water bath and positioned on the tip of the terminal leaflet of leaf 1 of seedlings at the two-leaf stage. After treatment the tomato seedlings were incubated at 30 °C for 48 h under fluorescent lights providing $200 \mu\text{mol m}^{-2} \text{ s}^{-1}$ of photosynthetically active radiation. After incubation the terminal leaflet of leaf 1 or leaf 2, as required, was harvested and immediately frozen in liquid N_2 . Proteinase inhibitor preparations were made by extracting leaf tissue with TES buffer (*n*-Tris(hydroxymethyl)methyl-2-aminoethane-sulphonic acid, 100 mM, pH 7.8) at 4 °C in proportions of 100 mg fresh weight to 1 ml TES. Chymotrypsin activity was determined spectrophotometrically by conversion of *N*-benzoyl-L-tyrosine *p*-nitroanilide to *p*-nitroaniline. Stock chymotrypsin solution (Sigma; product code C4129) was made up in HCl (25 mmol dm^{-3}) and the concentration adjusted with TES (100 mM, pH 7.8) to produce a rate of change of absorbance of 0.025 units min^{-1} at 405 nm and 20 °C in a control assay consisting of chymotrypsin (7–8 μg in 0.6 ml); 0.3 ml TES (100 mM, pH 7.8); and 0.3 ml of a solution of substrate (1.2 mM) in 50% dimethylsulphoxide/TES. In assays of proteinase inhibitor, plant extract (0.3 ml) was preincubated with the chymotrypsin solution (0.6 ml) at room temperature for 30 min before addition of substrate.

* Chymotrypsin activity is given as per cent of activity in non-plant controls.

by diffusive porometry (Automatic Porometer Mk 3, Delta-T Devices) before and after wounding. Stomatal resistance increased by $\sim 30\%$ in the 5 min after mechanical wounding of the cotyledon, and by $\sim 80\%$ in the 10 min after a heat stimulus to leaf 1. However, these resistances, either before or after wounding, are less than 20% of the resistance of closed stomata. These data confirm that the direction of transpiration is not reversed during the wound response and preclude movement of a chemical signal out of the wounded tissue via the xylem. Transport through the symplast is also a possibility for the movement of a chemical signal, but the speed of this process¹⁹, up to $15 \mu\text{m s}^{-1}$ in higher plants, is much too slow to account for our observations.

There are two leading possibilities for a physical signal: a hydraulic signal with secondary electrical consequences, or a propagated electrical signal (action potential). There is evidence^{20,21} in wheat leaves for a hydraulic signal that propagates at a speed of at least 100 mm s^{-1} from the site of wounding and is associated with electrical events recorded at the plant surface. A hydraulic pressure wave would propagate rapidly through the plant, possibly at speeds as high as that of sound in water ($1,500 \text{ m s}^{-1}$). We detected no systemic *pin* activity when the wounded cotyledon was excised before the electrical signal had exited to the stem (Table 1, line 10). In these measure-

ments, the mean time from wound to excision was 28 s, which should be sufficient for passage of a hydraulic signal. We conclude that the systemic response is caused by an electrical signal propagating through the plant.

We believe that an appropriate model for the electrical signal conduction pathway in plants may be provided by the well established epithelial conduction system of animals. Epithelial conduction was first demonstrated¹⁰ using nerve-free and muscle-free epithelia of siphonophores and hydromedusae and has now been found in at least four animal phyla¹¹. This mechanism involves the cell-to-cell conduction of action potentials through gap junctions at speeds of 2 to 500 mm s^{-1} . Although plants lack any structures comparable to animal nerves, cells in a plant tissue are linked by plasmodesmata that have solute permeabilities and electrical conductivities nearly identical to those of gap junctions in animal tissues^{22,23}. Thus plant tissues consisting of cells linked by plasmodesmata are analogous, in terms of their ability to conduct electrical signals, to animal epithelia. An electrical mechanism for long-distance signalling in plants may be much more general than has previously been recognized. Recent work has shown that many plant species can transmit action potentials and variation potentials^{24–26}. Such propagated electrical responses can often be instigated by stimuli such as osmotic shock, heat shock, cold and touch, which also

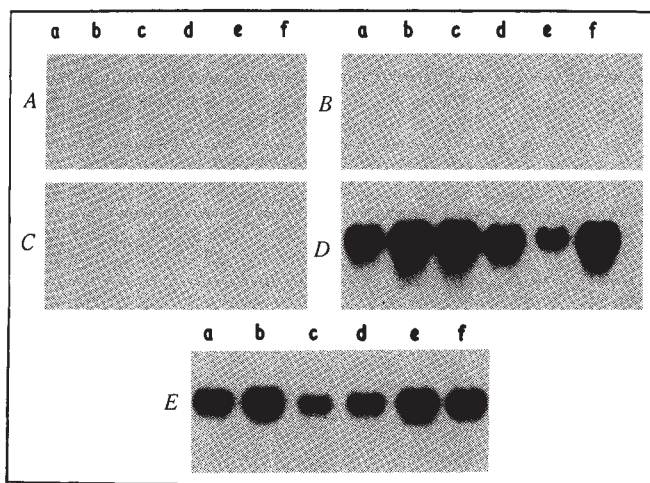


FIG. 3 Five conditions were applied to the tomato seedlings. A, No treatment; B, the petiole of the cotyledon was chilled as described in Fig. 1 legend, but no further treatment was applied; C, the petiole of the cotyledon was chilled as in B, and the cotyledon was then excised; D, the petiole of the cotyledon was chilled as in B, a mechanical wound was applied to the distal end of the cotyledon (see Table 1 legend), and the cotyledon excised 5 min after wounding; E, the petiole of the cotyledon was chilled as in B, a mechanical wound was applied as in D, but the cotyledon was not then excised. Following these conditions, seedlings were incubated at 30 °C for 4 h under fluorescent lights providing 200 $\mu\text{mol m}^{-2} \text{s}^{-1}$ photosynthetically active radiation before the terminal leaflet of leaf 1 was harvested and immediately frozen in liquid N_2 . RNA was extracted and analysed by northern blotting²⁷ using a ^{32}P -labelled complementary DNA probe to tomato *pin 2* (ref. 28). Equivalence of RNA loading in each track was confirmed spectrophotometrically. Each track is the analysis of 10 μg leaf RNA from a single plant that had also been analysed (conditions B–E) for electrical activity. A systemic electrical signal was recorded with all of the plants shown in D and E; no systemic electrical signals were observed with any of the plants shown in B and C.

elicit stress responses in plants²⁶.

To our knowledge, our results provide the first definite evidence for a link between an electrical signal and a biochemical response in plants. We conclude that an electrical signal is the underlying mechanism of the systemic wound response in our experimental system and are now investigating the relationship between the electrical signal and the wider range of chemical agents known to lead to *pin* induction. □

Received 11 June; accepted 15 September 1992.

- Ryan, C. A. *A. Rev. Phytopathol.* **28**, 425–449 (1990).
- Bowles, D. J. *A. Rev. Biochem.* **59**, 873–907 (1990).
- Ryan, C. A. & Farmer, E. E. *A. Rev. Pl. Physiol. Pl. molec. Biol.* **42**, 651–674 (1991).
- Pena-Cortes, H., Willmitzer, L. & Sanchez-Serrano, J. *Pl. Cell* **3**, 963–972 (1991).
- Pearce, G., Strydom, D., Johnson, S. & Ryan, C. A. *Science* **253**, 895–898 (1991).
- Farmer, E. E. & Ryan, C. A. *Pl. Cell* **4**, 129–134 (1992).
- Doherty, H. M., Selvendran, R. A. & Bowles, D. J. *Physiol. molec. Pl. Path.* **33**, 377–384 (1988).
- Doherty, H. M. & Bowles, D. J. *Pl. Cell Envir.* **13**, 851–855 (1990).
- Wildon, D. C., Doherty, H. M., Eagles, G., Bowles, D. J. & Thain, J. F. *Ann. Bot.* **64**, 691–695 (1989).
- Mackie, G. O. *Am. Zool.* **5**, 439–453 (1965).
- Anderson, P. A. V. *Prog. Neurobiol.* **15**, 161–203 (1980).
- Gasser, H. S. *Am. J. Physiol.* **97**, 254–270 (1931).
- Blatt, F. J. *Biochim. biophys. Acta* **339**, 382–389 (1974).
- Lang, A. & Minchin, P. E. H. *J. exp. Bot.* **37**, 389–398 (1986).
- Baker, D. A. & Milburn, J. A. (eds) *Transport of Photoassimilates* (Longman, Harlow, 1989).
- Pena-Cortes, H., Sanchez-Serrano, J., Rocha-Sosa, M. & Willmitzer, L. *Planta* **174**, 84–89 (1988).
- Farmer, E. E., Johnson, R. R. & Ryan, C. A. *Pl. Physiol.* **98**, 995–1002 (1992).
- Nelson, C. E., Walker-Simmons, M., Makus, D., Zuroske, G., Graham, J. & Ryan, C. A. in *Mechanisms of Plant Resistance to Insects* (ed. Hedin, P.) 103–122 (American Chemical Society, Washington DC, 1983).
- Tyree, M. T. *J. theor. Biol.* **26**, 181–214 (1970).
- Malone, M. & Stankovic, B. *Pl. Cell Envir.* **14**, 431–436 (1991).
- Malone, M. *Planta* **187**, 505–510 (1992).
- Robards, A. W. & Lucas, W. J. *A. Rev. Pl. Physiol. Pl. molec. Biol.* **41**, 369–419 (1990).
- Robards, A. W., Lucas, W. J., Pitts, J. D., Jongsma, H. J. & Spray, D. C. (eds) *Parallels in Cell to Cell Junctions in Plants and Animals* (Springer, Berlin, 1990).
- Pickard, B. *Bot. Rev.* **39**, 172–201 (1973).
- Pickard, B. *Naturwissenschaften* **61**, 60–64 (1974).
- Davies, E. *Pl. Cell Envir.* **10**, 623–631 (1987).
- Smallwood, M., Gurr, S. J., McPherson, M. J. & Bowles, D. J. *Biochem. J.* **281**, 501–505 (1992).
- Longmann, J., Schell, J. & Willmitzer, L. *Analyt. Biochem.* **163**, 16–20 (1987).

ACKNOWLEDGEMENTS. This work was supported by the AFRC. P.O.D. was supported by a studentship from the SERC.

Absence of aluminium in neuritic plaque cores in Alzheimer's disease

J. P. Landsberg, B. McDonald* & F. Watt†

Scanning Proton Microprobe Unit, Nuclear Physics Laboratory, University of Oxford, Oxford OX1 3RH, UK

* Department of Neuropathology, Radcliffe Infirmary, Woodstock Road, Oxford OX2 6HE, UK

CONTROVERSY exists over whether aluminium has a role in the aetiology of Alzheimer's disease. Alzheimer's disease is neuropathologically characterized by the occurrence of a minimum density of neurofibrillary tangles and neuritic plaques in the hippocampus and the association cortex of the brain^{1,2}. The purported association of aluminium with Alzheimer's disease is based on: (1) the experimental induction of fibrillary changes in the neurons of animals by the injection of aluminium salts into brain tissue^{3,4}; (2) reported detection of aluminium in neuritic plaques^{5–8} and tangle-bearing neurons^{9,10}; (3) epidemiological studies linking aluminium levels in the environment, notably water supplies, with an increased prevalence of dementia^{11–14}; and (4) a reported decrease in the rate of disease progression following the administration of desferrioxamine, an aluminium chelator, to clinically diagnosed sufferers of Alzheimer's disease¹⁵. Here we use nuclear microscopy, a new analytical technique involving million-volt nuclear particles, to identify and analyse plaques in post-mortem tissue from patients with Alzheimer's disease without using chemical staining techniques and fail to demonstrate the presence of aluminium in plaque cores in untreated tissue.

The most direct evidence for the link between aluminium and Alzheimer's disease (AD) stems from reports of granular aluminosilicates found in neuritic plaque cores. Neuritic plaques are extracellular structures measuring up to 200 μm in diameter, and classical core-containing neuritic plaques are composed of amyloid cores surrounded by a halo of abnormal neurites (Fig. 1). Several authors report the presence of aluminium^{5–8} and silicon^{7,16} in the plaque cores, although others^{17–19} have been unable to detect any increase in the level of these elements in neuritic plaques. Duckett and Galle^{5,6} using energy dispersive electron probe microanalysis, a technique with an analytical sensitivity of between 100 and 1,000 p.p.m., have reported finding aluminium in some plaque cores using fixed embedded tissue; Candy *et al.* used the electron probe on silver-stained tissue²⁰ and other microanalytical techniques^{7,21} and reported that most neuritic plaque cores contained aluminium co-localized with silicon^{7,22}. Studies by others on stained tissue using the electron probe^{17,19} and the proton probe¹⁸, which is more than an order of magnitude more sensitive, have failed to confirm these findings. Stern *et al.*²³, using toluidine blue-stained tissue also failed to find aluminium or silicon in the plaque cores, despite using a technique (laser microprobe mass analysis) capable of parts-per-million sensitivity, although they consistently find aluminium in neurofibrillary tangles using this technique.

We describe the application of nuclear microscopy to the analysis of *in situ* neuritic plaque cores using both stained and unstained AD tissue. Nuclear microscopy uses a focused MeV ion beam (for example, 3 MeV protons) and employs a group of analytical techniques that can be simultaneously applied to a sample^{24,25}. These include particle-induced X-ray emission (PIXE) which provides quantitative analytical information down to the parts-per-million level for elements above sodium in the periodic table by measuring the characteristic X-ray

† Present address: Department of Physics, University of Singapore, Blk S12, Lower Kent Ridge Road, 0511 Singapore.

TABLE 1 Results of scans of stained tissue

Case*	Region†	Total no. plaques scanned	Total no. Al deposits in scans	Number of plaque cores scanned	Number of plaque cores with Al	Number of background scans	Total no. Al deposits in backgrounds
A	tc	26	7	14	1	5	1
D	h	10	3	4	0	6	2
D	tc	6	2	5	1	3	1
E	h	9	2	6	0	7	1
E	tc	11	3	6	1	7	2
F	h	11	3	7	1	7	5
F	tc	10	2	6	0	7	3
G	h	12	0	9	0	6	1
G	tc	8	2	3	1	7	2
CB	h	2	0	1	0	12	2
CB	tc	0	0	0	0	14	4

The average scanned area was $75 \times 75 \mu\text{m}$. Eight per cent of the plaque cores were found to have accumulations of aluminium.

* Cases A to G refer to neuropathologically determined AD tissue, and case CB refers to control tissue.

† tc, Temporal cortex; h, hippocampus.

emission following proton/electron collisions²⁶; Rutherford backscattering spectrometry (RBS), which when applied to thin biological specimens, provides information on the matrix elements C, N and O by measuring the energy of the backscattered incident ion after a nuclear collision²⁷; and scanning transmission ion microscopy (STIM) which gives information on the structure and density of the sample by monitoring the energy loss of the transmitted ions^{28,29}.

Table 1 gives the results of the analysis of 105 neuritic plaques in stained tissue from the temporal cortex and hippocampus of five clinically and pathologically characterized cases of AD and two aged controls, and shows that less than 10% of plaque cores contain aluminium deposits, which are usually co-localized with silicon. Aluminium and silicon were also detected in the background and in the control tissue, and indicate that there is a possibility of contamination.

A significant fraction of airborne dust is composed of aluminosilicates up to about $4 \mu\text{m}$ in size³⁰, and these elements are ubiquitous as contaminants. PIXE analyses of the reagents (many of them ARISTAR quality) used in both immunohistochemical staining and silver-staining procedures showed trace levels of contaminants, despite filtering through $0.2\text{-}\mu\text{m}$ filters. Scans of residues from evaporated filtrates showed accumulations of inorganic elements (including aluminosilicates) up to

$4 \mu\text{m}$ across. These analyses have shown that many commercial reagents, including embedding medium, contain aluminosilicate contaminants in low concentrations ($< \text{p.p.m.}$), present either as a fine suspension of particles or in solution, when they crystallize out as the solvents evaporate; they may also preferentially locate on the sulphate and phosphate moieties present in the plaques.

Nuclear microscopical investigations using unstained tissue were carried out on more than 80 plaque cores from the temporal cortex and hippocampus of four cases of AD and two aged controls. The plaques were located using STIM imaging and unambiguously identified using the following criteria. (1) The shape of the core and associated corona from the STIM image is consistent with the optical images in stained sections; (2) the core is more dense than the corona, which in turn is more dense than the surrounding tissue; (3) the elemental distribution of the plaque as measured by PIXE and RBS shows similar minor and major elemental composition to those analysed in stained plaques, that is, increased levels of elemental sulphur, phosphorus and nitrogen; and (4) examination of stained serial sections shows similar plaque density per area of tissue, and the plaques are located in a similar position with respect to more obvious common features such as blood vessels. As a control, nearly 30 mm^2 of plaque-free tissue was scanned and

TABLE 2 Results of scans of unstained tissue

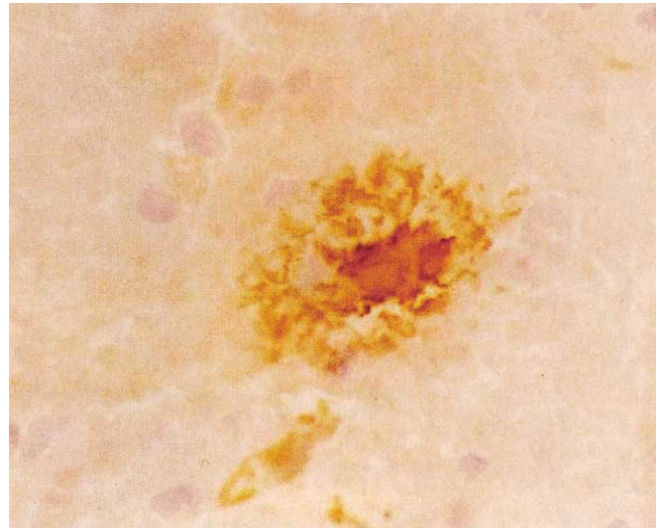
Case*	Region†	Total no. plaques scanned	Total no. Al deposits in scans	Number of plaque cores scanned	Number of plaque cores with Al	Number of background scans	Total no. Al deposits in backgrounds
B	h	10	0	10	0	5	0
B	tc	10	1	10	0	7	0
C	h	9	0	9	0	13	0
C	tc	9	2	9	0	3	0
D	h	11	1	11	0	12	1
D	tc	12	1	12	0	16	2
E	h	6	0	6	0	6	1
E	tc	11	0	11	0	13	0
CA	tc					10	2
CA	h					10	1
CB	tc					10	1
CB	h					13	0
CC	tc					9	2
CC	h					10	0

The average area scanned was $100 \times 100 \mu\text{m}$. Aluminium was not found in any plaque core. The small amounts of aluminium focal deposits found in the background were caused by contamination in the photoform film supporting the tissue.

* Cases CA, CB and CC refer to control (non-AD) tissue.

† tc, Temporal cortex; h, hippocampus.

FIG. 1 Optical micrograph of immunohistochemically stained core containing neuritic plaque. Tissue blocks were dissected at post-mortem from the hippocampus (at the level of the lateral geniculate body) and the middle temporal gyrus (Brodmann's area 21) of both the AD and control cases. Brain tissue was placed on cryostat cork discs using a small amount of embedding medium (OCT compound, Miles). The tissue was then snap-frozen by immersion in liquid nitrogen before storage at -70°C for up to six months. Sections ($10\text{ }\mu\text{m}$) were cut using a cryostat microtome, transferred onto $0.5\text{-}\mu\text{m}$ polyolefin-coated cleaned glass slides and stored at -20°C . Immunohistochemical staining using antiserum reacting against complement factor C3d was used to visualize neuritic plaques. This organic staining method was chosen to minimize the risk of contamination by inorganic elements. Double-distilled water (milli-Q quality) and ARISTAR grade reagents were used throughout; all immunostaining procedures were undertaken in a closed container to minimize airborne contamination. Brain sections were incubated for 1 h with rabbit polyclonal C3d antiserum (Dakopatts) at a dilution of 1 in 100 in 50 mM Tris-HCl, pH 7.4, containing 200 mM NaCl in a closed chamber at room temperature, followed by three 5-min washes in the same Tris buffer. The sections were then overlaid with peroxidase-conjugated swine anti-rabbit antiserum (Dakopatts) at a dilution of 1 in 100 in Tris buffer for 1 h at room temperature, washed three times with Tris buffer and the bound antibody complex visualized using diaminobenzidine (Sigma; 60 mg per 100 ml in Tris containing $40\text{ }\mu\text{l}$ 30% v/v H_2O_2). Substrate was applied to the sections and incubated for 3–5 min, resulting in a brown reaction product. The reaction was terminated by washing the sections in milli-Q water.



although numerous corpora amylacea were identified, no features were observed that fitted the plaque identification criteria.

Figure 2 shows a typical STIM image of a plaque in unstained tissue, together with the corresponding elemental maps of S, P, C and N. The results of analyses on unstained tissue is shown in Table 2, which indicates that aluminium was not observed in any of the neuritic plaque cores scanned, at a sensitivity of

15 p.p.m. This detection limit was determined from analysis of aluminium standards made from solutions of aluminium nitrate in poly(ethyl methacrylate) resin.

Because the analysis of unstained AD tissue using the sensitive techniques of nuclear microscopy does not indicate increased levels of aluminium in neuritic plaque cores, we now believe that previous evidence that aluminium is involved in the aetiology of Alzheimer's disease should be reviewed to take into

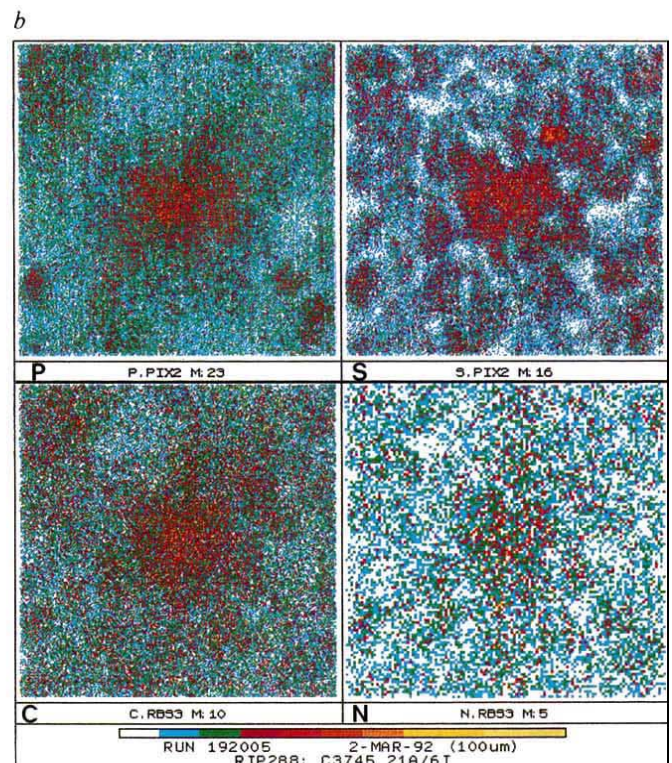
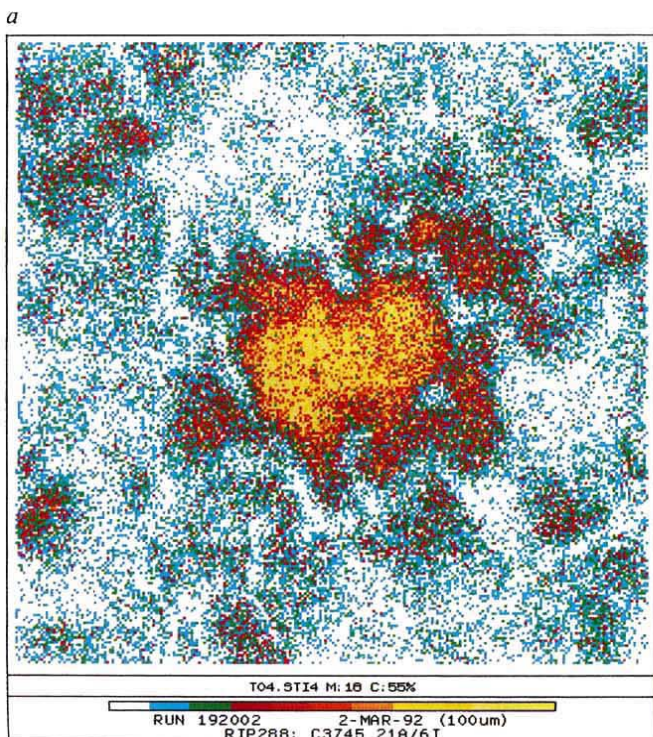


FIG. 2 *a*, STIM image of a core-containing neuritic plaque from freeze-dried unstained tissue. The map is shown in false colour, where yellow represents the most dense and blue the least dense areas. *b*, Elemental maps of the core-containing plaque shown in *a*; the sulphur (S) and phosphorus (P) maps were obtained using PIXE and the carbon (C) and nitrogen (N) maps using

RBS. The $100\times 100\text{ }\mu\text{m}$ maps are shown in false colour, with yellow representing the highest elemental concentration and blue the lowest. The maps were obtained using the Oxford Scanning Proton Microprobe operating with 3-MeV protons focused to a $1\text{-}\mu\text{m}$ spot.

account the probable contamination of tissue by aluminosilicates present in most reagents. □

Received 5 May; accepted 11 September 1992.

1. Khachaturian, Z. S. *Arch. Neurol.* **42**, 1097–1105 (1985).
2. Mirra, S. S. *et al. Neurology* **4**, 479–486 (1991).
3. Terry, R. D. & Peña, C. J. *Neuropath. exp. Neurol.* **24**, 200–210 (1965).
4. Klatzo, I., Wisniewski, H. & Streicher, E. J. *Neuropathol. exp. Neurol.* **24**, 187–199 (1965).
5. Duckett, S. & Galle, P. C. *Acad. Sci. D* **282**, 393–395 (1976).
6. Duckett, S. & Galle, P. J. *Neuropath. exp. Neurol.* **39**, 350 (1980).
7. Candy, J. M. *et al. Lancet* **i**, 354–357 (1986).
8. Masters, C. L. *et al. EMBO J.* **4**, 2757–2763 (1985).
9. Perl, D. P. & Brody, A. R. *Science* **208**, 297–299 (1980).
10. Good, P. F., Perl, D. P., Bierer, L. M. & Schmeidler, J. *Ann. Neurol.* **31**, 286–292 (1992).
11. Martyn, C. N. *et al. Lancet* **i**, 59–62 (1989).
12. Flaten, T. P. *Environ. Geochem. Health* **12**, 152–167 (1990).
13. Neri, L. C. & Hewitt, D. *Lancet* **338**, 390 (1991).
14. Michel, P. *et al. in Alzheimer's Disease: Basic Mechanisms, Diagnosis and Therapeutic Strategies* (eds Iqbal, K., McLachlan, D. R. C., Winblad, B. & Wisniewski, H. M.) 387–391 (Wiley, New York, 1991).
15. Crapper, D. R. *et al. Lancet* **337**, 1304–1308 (1991).
16. Nikaido, T., Austin, J., Trueb, L. & Rivehart, R. *Arch. Neurol.* **27**, 549–554 (1972).
17. Jacobs, R. W. *et al. J. Neurol. Sci.* **16**, 498–503 (1989).
18. Landsberg, J. P. *et al. Nucl. Instr. Meth.* **B54**, 180–185 (1991).
19. Chafi, A. H. *et al. Neurosci. Lett.* **123**, 61–64 (1991).
20. Candy, J. M. *et al. in Senile Dementia of the Alzheimer Type* (eds Traber, J. & Gispen, W. H.) 183–197 (Springer, Berlin, Heidelberg, 1985).
21. Candy, J. *et al. in Interdisciplinary Topics in Gerontology* Vol. 25, 140–155 (Karger, Basel, 1988).
22. Edwardson, J. A. & Candy, J. M. *Environ. Geochem. Health* **12**, 94–96 (1990).
23. Stern, A. J. *et al. J. Neuropath. exp. Neurol.* **45**, 361 (1986).
24. Watt, F., Landsberg, J. P. & Grime, G. W. *X-ray Microanalysis in Biology: Experimental Techniques and Applications* (eds Sigee, D. C., Summer, A. T. & Warley, A.) (Cambridge University Press, UK, 1992).
25. Watt, F. & Grime, G. W. (eds) *Principles and Applications of High-Energy Ion Microbeams* (Hilger, Bristol, 1987).
26. Johansson, S. A. E. & Campbell, J. L. *PIXE, A Novel Technique for Elemental Analysis* (Wiley, Chichester, 1988).
27. Chu, W.-K., Mayer, J. W. & Nicolet, M. A. *Backscattering Spectrometry* (Academic, New York, 1978).
28. Schofield, R. M. S. *et al. Nucl. Instr. Meth.* **B30**, 398–403 (1988).
29. Bench, G. S. & Legge, G. J. F. *Nucl. Instr. Meth.* **B40/A1**, 655–658 (1989).
30. Artaxo, P., Andrade, F. & Maenhaut, W. *Nucl. Instr. Meth.* **B49**, 383–387 (1989).

ACKNOWLEDGEMENTS. This research was funded by the Wellcome Trust.

Real and optimal neural images in early vision

J. H. van Hateren

Department of Biophysics, University of Groningen, Westersingel 34, NL-9718 CM Groningen, The Netherlands

IT has been suggested^{1–3} that the first steps in visual processing strive to compress as much information as possible about the outside world into the limited dynamic range of the visual channels. Here I compare measured neural images with theoretical calculations based on maximizing information, taking into account the statistical structure of natural images. Neural images were obtained by scanning an image while recording from a second-order neuron in the fly visual system. Over a 5.5-log-units-wide range of mean intensities, experiment and theory correspond well. At high mean intensities, redundancy in the image is reduced by spatial and temporal antagonism. At low mean intensities, spatial and temporal low-pass filtering combat noise and increase signal reliability.

Barlow¹ argued that an important task of early vision is to reduce the redundancy of natural images. Reducing the redundancy may be useful because it potentially increases the amount of information that is transferred through the channels of the visual system^{4,5}. Natural images are redundant because their intensity values do not vary randomly from point to point, but are correlated^{2,6}. Correlation occurs, for instance, across the surface of objects. Apart from this spatial correlation, images are also correlated in time. Reducing these redundancies will result in spatial and temporal sharpening of the image, thus enhancing edges and temporal change. It leads to phenomena such as lateral inhibition^{1,2} and self-inhibition².

Although redundancy reduction may be a good strategy when the image is of high quality, such as under bright light conditions, it becomes counterproductive at low light levels⁴. Then it tends to enhance spatial and temporal fluctuations mainly due to noise. Indeed, requiring the system to adapt to different signal-to-noise ratios to maximize information results in reduced image redundancy only when the signal-to-noise ratio is high⁷. At low signal-to-noise ratios, image redundancy may even be increased, leading to phenomena such as spatial pooling^{5,8} and increased temporal integration⁷.

A theory that enables quantitative predictions of the optimal spatial and temporal filtering at the various signal-to-noise ratios is illustrated in Fig. 1a. The stimulus reaches a visual channel, one of an array of similar channels that collect the image. The stimulus is first transduced by a prefilter which is noisy itself, and is low-pass in both space (because of optical limitations of the lens and photoreceptors) and time (because of the limited speed of phototransduction). Noise at the prefilter originates both from the stimulus (photon noise) and from the photoreceptor itself. The signal is, by means of a neural filter, delivered to a noisy information channel supporting only a limited range of response values. Thus the channel has a limited dynamic range and information capacity⁹. The essence of the theory is that it requires the neural filter to shape the signal such that it fits the

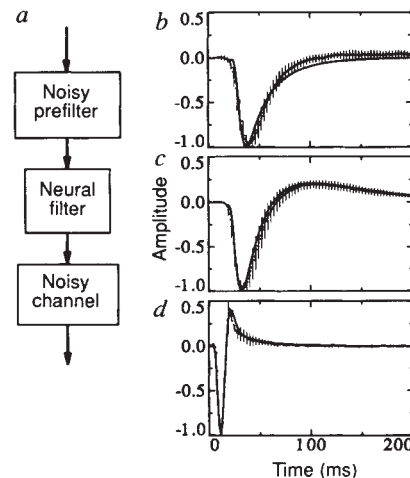


FIG. 1 a. Outline of the theory. An image is filtered in space and time by a noisy prefilter. A neural filter then transfers the image to a noisy channel of limited dynamic range. The neural filter maximizes the amount of information in the channel, while keeping the response range within the limits of the channel. Let S , N_p , and N_c be the power spectra of stimulus, prefilter noise, and channel noise, respectively, and p_p and p_n the power transfer functions of the prefilter and neural filter. Then the information rate in the channel⁹ is $R = \int \log(1 + \text{signal power}/(\text{noise power})) df = \int \log(1 + S p_p p_n / (N_p p_n + N_c)) df$, with integration over temporal and spatial frequencies. The required response range $K = (\int S p_p p_n + N_p p_n + N_c df)^{1/2}$. Maximizing the information rate R while keeping K fixed to the channel's dynamic range is solved using a Lagrange multiplier⁹, yielding an expression for p_n (ref. 7). For related approaches in the spatial domain see refs 5, 8, and 10. b–d. Responses to short flashes (much shorter than the photoreceptor's integration time) measured in fly LMCs (dots) and theoretical calculations (continuous lines). Figures adapted from ref. 7. b. Response to a point source, centred at the receptive field, and superimposed on a background of 10 effective photons per second per photoreceptor. Dots and error bars show mean and standard deviations of measurements in 8 LMCs, each averaged over at least 320 stimulus presentations and subsequently normalized. Responses were within the linear range of the cells. c. Response to a uniform wide field superimposed on the same background as in b. Mean of 8 LMCs. d. As c, background 3×10^6 photons per second per photoreceptor. Mean of 6 LMCs. The theoretical calculations are based on an average contrast of natural images of 0.4 (ref. 11), a $1/f_s^2$ -behaviour of the spatial power spectrum^{12,13}, with f_s the spatial frequency, and the temporal part of the power spectrum caused by a distribution of image velocities mainly due to movements of the animal itself⁷.

channel's dynamic range and feeds the channel with the maximum possible amount of information about the stimulus. Given the power spectra of stimulus, prefilter noise and channel noise, and given the channel's dynamic range, it is possible to calculate the neural filter accomplishing this task⁷. This filter is unique for a given level of prefilter noise, and thus for a given light intensity.

As a test of the predictive power of the theory, it was applied to the responses of second-order neurons in the lamina neuropile

in the blowfly visual system⁷. These large monopolar cells (LMCs) are directly postsynaptic to the photoreceptors, hyperpolarize in response to intensity increments and depolarize to decrements, and transport their information by means of graded potentials to the next neuropile. Figure 1*b-d* shows examples of experimental and theoretical responses to short light flashes. Experimental and theoretical responses behave similarly: both are less transient for point sources (Fig. 1*b*) than for wide-field stimuli (Fig. 1*c*), and become much faster and display increased spatial and temporal antagonism at higher mean intensities (Fig. 1*d*).

How will the spatial and temporal filtering done by the LMCs affect the image? To construct the resulting neural image, one would ideally have to measure the responses of the entire array of LMCs looking at an image changing in time. This is clearly not possible with present techniques, but it is possible to mimic this by recording the response of a single LMC while it is scanning an image. To this end an image was moved horizontally through the LMC's receptive field at 32 consecutive vertical scan positions. This yielded 32 scan lines of the neural image, which were subsequently used to construct neural images as shown in the left column of Fig. 2. The right column of Fig. 2 shows calculated responses, produced by filtering the moving image with the spatiotemporal filters calculated from the theory explained above (Fig. 1*b-d*). Subsequent rows in Fig. 2 correspond to increasing mean intensities, spanning a total range of 5.5 log units.

At low mean intensities, the images are low-pass filtered versions of the stimulus, apparently to combat photon noise. At high mean intensities, spatial edges and temporal transients are enhanced, at the expense of areas of uniform intensity. Note, for instance, that the intensities of the white wall and the doorway are very different in the stimulus (Fig. 2*a*), but hardly differ in the response (Fig. 2*m*).

As Fig. 2 shows, the neural image in the LMC array is very well predicted by a theory based on a remarkably simple

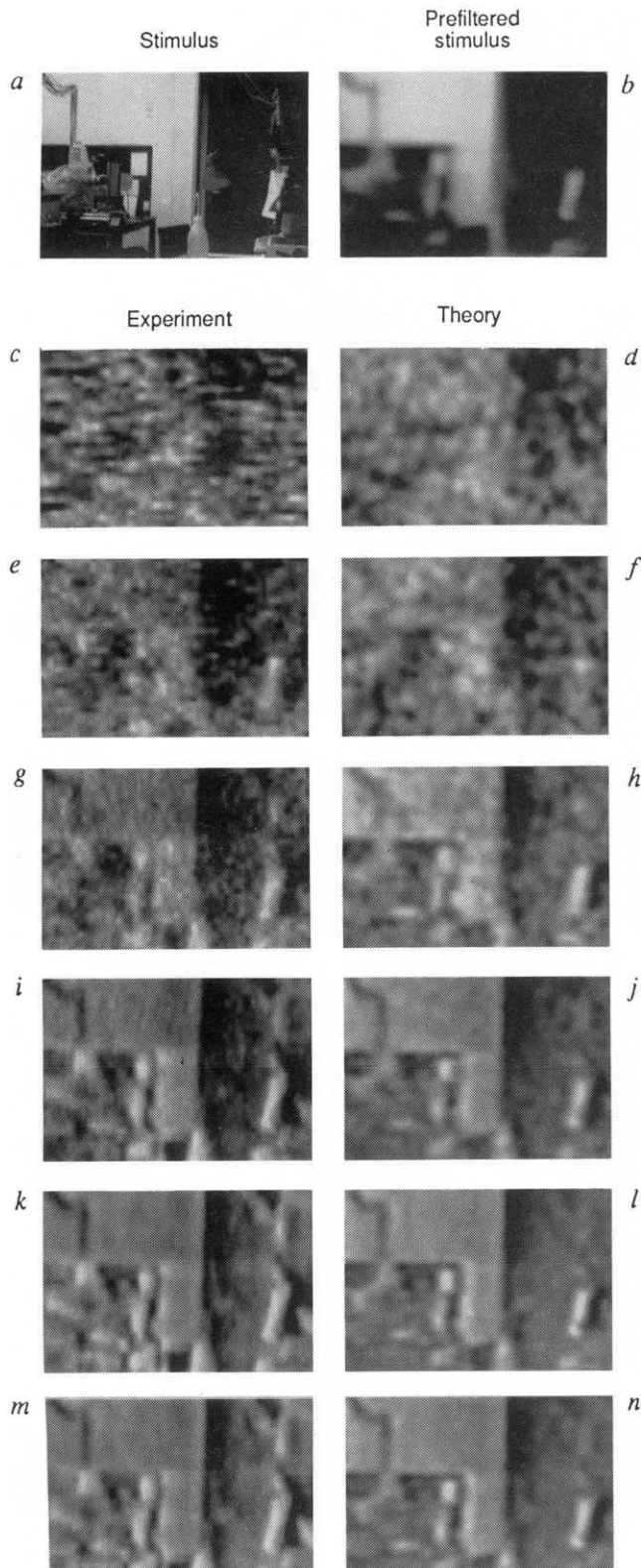


FIG. 2 Measurements and theoretical calculations of neural images. *a*, Stimulus scanned by the LMC. *b*, Prefiltered stimulus. The prefilter has fixed spatial and temporal properties, and was determined from the limits of photoreceptor performance, that is, when fully light-adapted. It consists of a gaussian spatial filter with a standard deviation $\sigma_s = 0.268$ cycles per deg, and a negative exponential temporal filter with a $1/e$ value $\sigma_t = 45$ Hz (ref. 7). Measurements (left column) were performed at mean intensities (in number of effective photon absorptions per second per photoreceptor): *c*, 10 (about the light-level at night); *e*, 1×10^2 (street light); *g*, 2×10^3 (dim interior light); *i*, 2×10^4 (interior light); *k*, 2×10^5 (daylight, overcast sky); and *m*, 3×10^6 (sunny day). The corresponding theoretical predictions (right column) have signal-to-noise ratios at the prefilter of: *d*, 0.5; *f*, 1.5; *h*, 4.5; *j*, 11; *l*, 22; and *n*, 21, estimated from measurements in the photoreceptors⁷. METHODS. Intracellular recordings were made from fly LMCs by standard methods¹⁴. The stimulus was selected to have a reasonably wide variety of features, and had a contrast and spatial frequency spectrum falling within the normal range of natural images⁷. It was mounted on an xy-recorder, and moved under computer control. Illumination and calibration were as described in ref. 7. The stimulus, measuring $58^\circ \times 41^\circ$, was scanned in 32 rows at a speed of 110 deg s^{-1} ; the figure shows results for the cell scanning the stimulus (*a*) from left to right. The speed of 110 deg s^{-1} was chosen for technical reasons, but is well within the range of speeds encountered by the normally behaving fly. It is the fly's equivalent of less than 1 deg s^{-1} for the human fovea, with its 50 times higher spatial and 4 times lower temporal resolution. The reconstructed images (left column) were mildly low-pass filtered vertically (spatial cut-off frequency twice that of the photoreceptor) in order to reduce high spatial frequency artefacts (still quite obvious in the noisy low-intensity measurements) due to the rows of the scanning method. Brighter pixels denote larger hyperpolarizations. The figures show only single sweeps (not averages). Complete sets of results (5-10 images at each mean intensity) were obtained in 18 LMCs, partial sets in another 11 cells. These images were, apart from fluctuations due to noise, virtually identical.

principle, namely maximization of information. But there are slight differences at high mean intensities; the measured images are generally somewhat sharper than the theoretical ones, have a better contrast, and also show a more completely suppressed response to wide areas of uniform intensity. This was consistently found in similar measurements of responses of 17 other LMCs. A likely explanation is that the theory assumes that an identical filter is applied to all parts of the image. A version of the theory allowing the filter to adapt locally, presently under development, lifts this restriction. □

Received 5 June; accepted 15 September 1992.

1. Barlow, H. B. in *Sensory Communication* (ed. Rosenblith, W. A.) 217–234 (MIT Press, Cambridge, MA, 1961).
2. Srinivasan, M. V., Laughlin, S. B. & Dubs, A. *Proc. R. Soc. Lond. B* **216**, 427–459 (1982).
3. Laughlin, S. B. *Trends Neurosci.* **10**, 478–483 (1987).
4. Kretzmer, E. R. *Bell Lab. Record Nov.*, 401–404 (1954).
5. Atick, J. J. & Redlich, A. N. *Neural Comput.* **2**, 308–320 (1990).
6. Kretzmer, E. R. *Bell Syst. tech. J.* **31**, 751–763 (1952).
7. van Hateren, J. H. *J. comp. Physiol. A* **171**, 157–170 (1992).
8. Tsukamoto, Y., Smith, R. G. & Sterling, P. *Proc. natn. Acad. Sci. U.S.A.* **87**, 1860–1864 (1990).
9. Goldman, S. *Information Theory* (Dover, New York, 1953).
10. Atick, J. J. *Network* **3**, 213–251 (1992).
11. Laughlin, S. B. in *Physical and Biological Processing of Images* (eds Braddick, O. J. & Sleight, A. C.) 42–53 (Springer, Berlin, 1983).
12. Burton, G. J. & Moorhead, I. R. *Appl. Opt.* **26**, 157–170 (1987).
13. Field, D. J. *J. opt. Soc. Am. A* **4**, 2379–2394 (1987).
14. Laughlin, S. B. & Hardie, R. C. *J. comp. Physiol.* **128**, 319–340 (1978).

ACKNOWLEDGEMENTS. I thank N. M. Jansoni and D. G. Stavenga for their comments. This research was supported by the Netherlands Organization for Scientific Research (NWO).

Presynaptic release probability influences the locus of long-term potentiation

Alan Larkman, Timo Hannay, Ken Stratford & Julian Jack

The University Laboratory of Physiology, Oxford University, Parks Road, Oxford OX1 3PT, UK

THE quantal hypothesis proposes that chemical synaptic transmission involves the probabilistic release of multimolecular packets of transmitter¹. Analysis of the resulting trial-to-trial fluctuations in postsynaptic response can provide estimates both of the number of quanta released and of the size of their postsynaptic effect. This in turn permits the quantification of the relative contributions of pre- and postsynaptic factors to the strength of a given synapse. Quantal analysis of excitatory synapses in the hippocampus has proved difficult^{2–6} and has led to contradictory conclusions when applied to long-term potentiation^{7–14}. Here we report the use of a combination of quantal analysis procedures to provide evidence that both pre- and postsynaptic changes can contribute substantially to the maintenance of long-term potentiation in the CA1 region of the hippocampus. The initial setting of the presynaptic release mechanism seems to determine their relative importance.

A powerful form of quantal analysis for central synapses is the study of amplitude histograms. The presence of equally spaced peaks in histograms is evidence of quantal behaviour, and the peak spacing gives a direct measure of the quantal size (the postsynaptic effect of a single quantum). Figure 1 shows an exemplar excitatory postsynaptic potential (e.p.s.p.) that was evoked for 250 trials at 0.1 Hz, and whose mean and standard deviation (s.d.) showed little change over time (Fig. 1b). The histogram for all the trials shows three clear peaks with a mean separation of 280 μ V (Fig. 1c). For this e.p.s.p. the signal-to-noise level was good (noise s.d. = 70 μ V) and spontaneous events could be resolved (Fig. 1d). The distribution of spontaneous amplitudes showed a major peak at 250 μ V, which is in

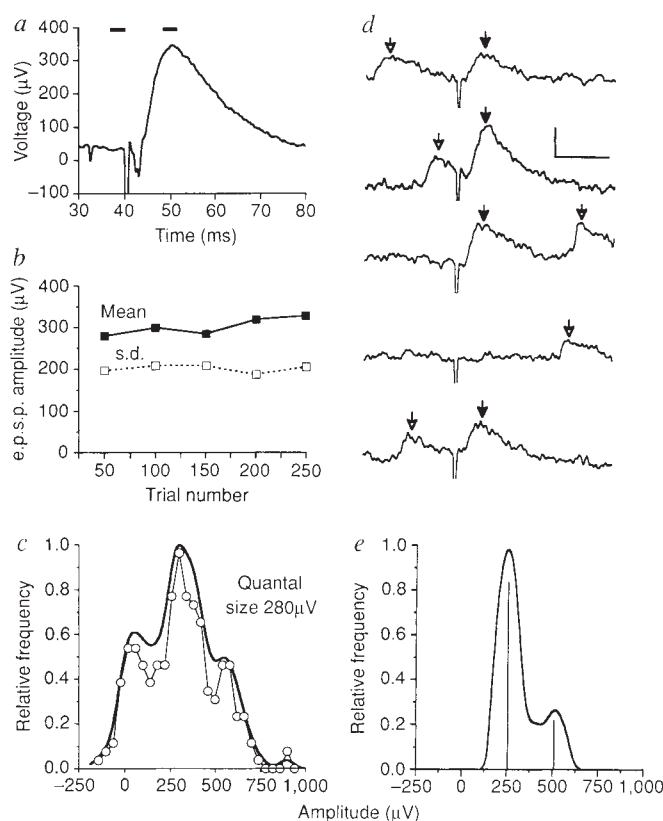
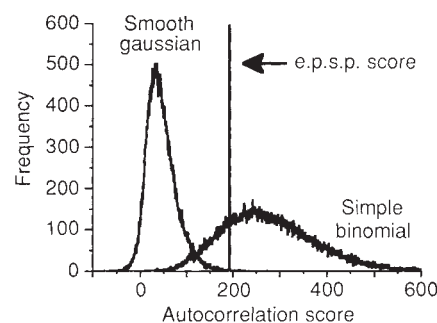


FIG. 1 Analysis of an e.p.s.p. for which 250 trials were evoked at constant 0.1 Hz. *a*, Averaged waveform of all 250 trials. Horizontal bars above waveform indicate zones used for measuring peak amplitudes on individual trials. The 10–90% rise time for this e.p.s.p. was 3.6 ms. *b*, The time course of e.p.s.p. peak amplitude mean (filled squares) and s.d. (open squares), binned into 5 epochs of 50 trials. The mean shows a slight increase during the recording period. *c*, Amplitude histogram for all 250 trials, binned finely and smoothed (thick line) and binned at 40 μ V but unsmoothed (open circles and thin line; relative frequency scale reduced slightly for clarity). The histogram shows three clear peaks with mean separation 280 μ V. The relative heights of the peaks are well described by a simple binomial distribution with $n=3$ and $p=0.34$. *d*, Individual records showing evoked (filled arrows) and spontaneous (open arrows) e.p.s.ps. Scale bars, 500 μ V and 20 ms. *e*, Smoothed histogram of the peak amplitudes of the 34 spontaneous events recorded for this e.p.s.p. The histogram was well fitted by the sum of two gaussians, each of s.d. 70 μ V, with their means and relative heights indicated by the vertical bars. Note that the major peak is at 250 μ V whereas the smaller peak is at about twice this amplitude. This is consistent with the quantal size of 280 μ V obtained from the histogram of evoked responses shown in *c*.

METHODS. Transverse hippocampal slices from young adult albino rats (Sprague-Dawley strain; 140–200 g) were maintained in an interface-type recording chamber at 34 °C. Conventional sharp-electrode voltage recordings were made from pyramidal neurons in the CA1 region. For most experiments, the Ca^{2+} and Mg^{2+} concentrations in the bathing medium were elevated to 4 mM, and picrotoxin (100 μ M) and glutamine (0.5 mM) were added. Some experiments were done using 2.5 mM Ca^{2+} and 2.5 mM Mg^{2+} . Small e.p.s.ps were evoked by bipolar wire electrodes placed in the stratum radiatum, with the stimulus strength adjusted to give the smallest e.p.s.p. that seemed to be evoked reliably. The e.p.s.ps were recorded on computer disk and their amplitudes measured using an automated computer routine that measured the average voltage difference between two zones, one in the pre-stimulus baseline period (avoiding the stimulus artefact if necessary) and the other straddling the peak of the averaged e.p.s.p. The positions and durations of these zones were chosen by inspection of the averaged e.p.s.p. waveform (see *a*). The noise amplitude distribution was obtained using zones of the same duration and separation, but both located in the pre-stimulus period. Amplitude histograms were routinely binned finely (5 μ V) and smoothed using a moving gaussian filter⁴ or by low-pass filtering in the Fourier domain. The locations of peaks in the amplitude histograms were determined visually, making use of an interactive computer routine.

FIG. 2 Illustration of the Monte Carlo computer simulations performed to assess the likelihood that peaky histograms could have arisen by sampling error from an underlying smooth distribution. The procedure made use of autocorrelation^{19,20} as an objective measure of the peakiness and equality of peak spacing in histograms. For a given histogram (in this case the histogram from Fig. 1c), an 'autocorrelation score' was calculated by subtracting the best-fitting smooth distribution (in this and many other examples a single gaussian provided a good fit), filtering the resulting difference function in the Fourier domain, then detecting peaks in its autocorrelation function. The height of the first peak in the autocorrelation function was taken as the autocorrelation score for that histogram. Amplitude histograms with evenly spaced peaks and a high 'peak-to-valley' ratio will give relatively high positive scores. The amplitude histogram shown in Fig. 1c scored 190 using this procedure. The left-hand histogram shows the distribution of autocorrelation scores of 50,000 samples, each of 250 trials, drawn from a smooth generator distribution, in this case a single gaussian of the same mean and s.d. as the e.p.s.p. histogram. The mean autocorrelation score for the random samples from this generator was 38 ± 28 , and only a very small proportion of samples (less than 1 in 1,000) scored as highly as the recorded e.p.s.p. histogram. In contrast, the right-hand histogram shows the



distribution of autocorrelation scores of 50,000 samples of 250 trials drawn from a quantal simple binomial generator similar to the data ($n=3$; $P=0.34$). The mean of this distribution is 270 ± 97 . Thus the data for this e.p.s.p. are very unlikely to have been generated from the smooth distribution, but are more typical of samples drawn from the quantal simple binomial generator.

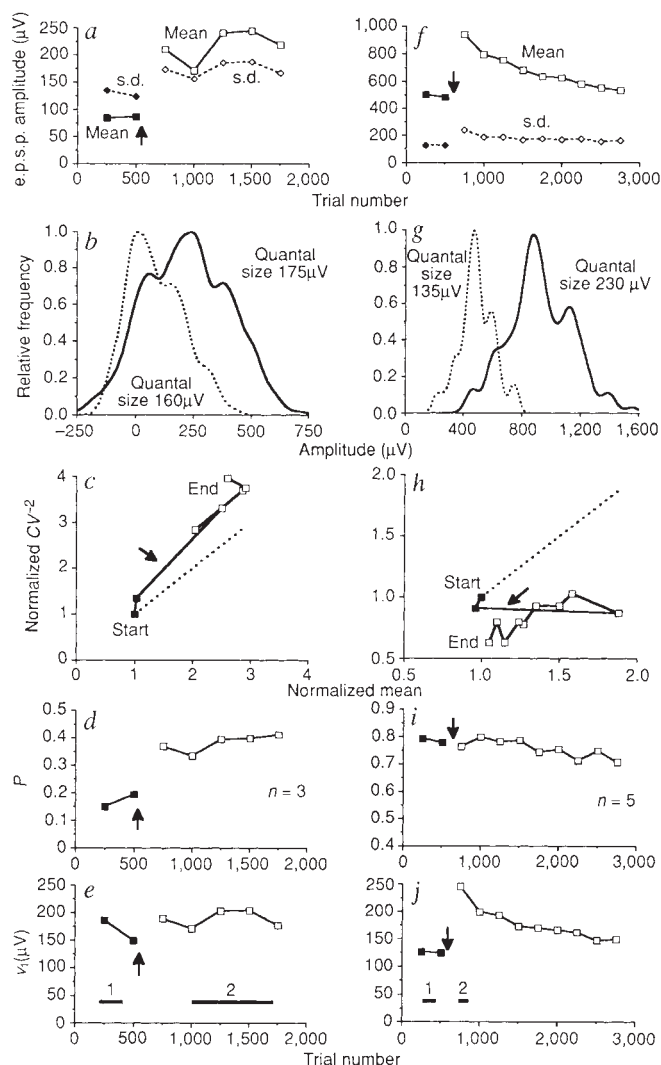
reasonable agreement with the quantal size of the evoked response. Monte Carlo simulations indicated that it was extremely unlikely that the peaks in the histogram of evoked responses had arisen by sampling artefact⁵ from a distribution that was smooth (Fig. 2).

Thus, in favourable cases, amplitude histograms can permit a reliable analysis of quantal size. If the quantal size changes over time, however, the histogram peaks become smeared and

the analysis can, at best, only be done for relatively brief selected periods. This was our experience for most hippocampal e.p.s.ps, even during stimulation at a constant rate⁴. We have therefore sought additional analysis strategies that could make use of all the recorded data. The time courses of changes in the mean and variance of the e.p.s.p. amplitude distribution are essentially independent of the occurrence or spacing of peaks in amplitude histograms, and can provide additional information about

FIG. 3 Analysis of two e.p.s.ps showing different forms of LTP following tetanic stimulation. Tetanus was given after trial 500 in both cases (arrowed). *a-e*, E.p.s.p. showing predominantly an increase in the probability of release. The e.p.s.p. mean (*a*, squares) shows a roughly two-fold increase after the tetanus. Amplitude histograms (*b*) taken from epochs before (dashed line; 200 trials) and after (solid line; 700 trials) the tetanus show a small increase in peak separation, but a major change in the relative heights of the peaks, with right-hand peaks relatively higher after the tetanus. The likelihood of these histograms having arisen by random sampling from smooth generator distributions was 1 in 35 before and 1 in 26 after the tetanus. The trajectory of the $1/CV^2$ plot (*c*) is steep and above the diagonal as the mean changes after the tetanus (arrowed). Using equations (1) and (2) with $n=3$ indicates that P (*d*) was initially low (<0.2) and roughly doubles, whereas v_1 shows only a minor increase (*e*). *f-j*, E.p.s.p. showing predominantly an increase in quantal size. The e.p.s.p. mean (*f*) again shows a roughly twofold increase after the tetanus, but then declines. Amplitude histograms (*g*) taken from epochs before (dashed line; 100 trials) and after (solid line; 80 trials) the tetanus show an increase in peak separation, but very little change in the relative heights of peaks. In spite of the small numbers of trials, the chances of these histograms having arisen from smooth distributions are less than 1 in 100 before and 1 in 60 after the tetanus. The $1/CV^2$ plot (*h*) shows a virtually horizontal trajectory following the tetanus. Equations (1) and (2) with $n=5$ indicate that P was initially high (0.8) and shows very little change (*i*), but v_1 almost doubles after the tetanus (*j*).

METHODS. A small e.p.s.p. was evoked at 1 Hz for usually 500 trials. There followed a pause of 1 min before the tetanic stimulation which consisted of five trains, each of 20 shocks at 100 Hz, repeated at 15 s intervals. The tetanus of the small e.p.s.p. was paired with the simultaneous tetanus of additional fibres at a similar depth in the stratum radiatum, together producing an e.p.s.p. usually large enough to cause the neuron to fire action potentials during each tetanus. There was then a further 1-min pause, after which 1-Hz test stimulation of the small e.p.s.p. was resumed. The success rate for obtaining an enhancement that persisted for more than 15 min was less than 10%. The mean 10–90% rise time for the 13 e.p.s.ps analysed was 2.8 ± 0.9 ms before LTP induction and 3.0 ± 1.1 ms after. Use of the e.p.s.p. mean and variance to estimate P and v_1 is invalid if the amplitude distribution is bimodal, as can be produced if there are failures of stimulation. Offsets, perhaps caused by extracellular field potentials, will, if uncorrected, also introduce errors. We have encountered both these difficulties with some hippocampal e.p.s.ps, but they were not serious for those analysed here.



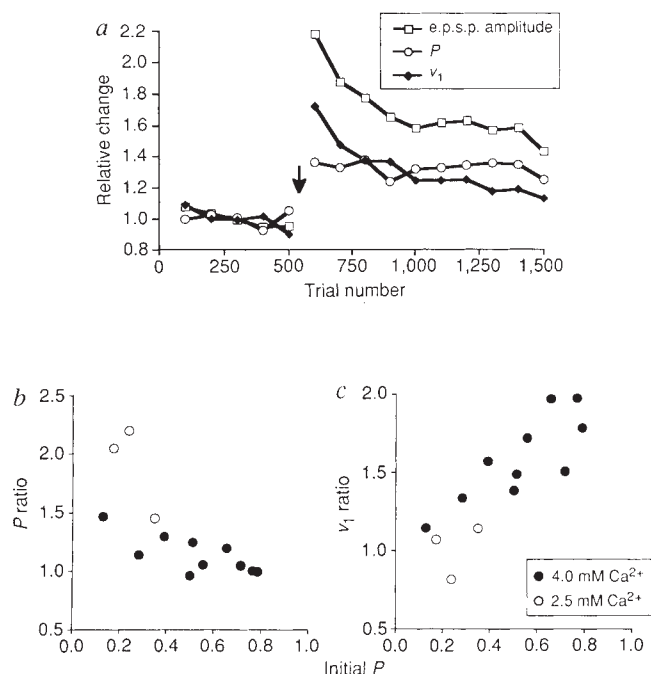


FIG. 4 *a*, Summary of the time course of the relative changes in e.p.s.p. peak amplitude (open squares), P (open circles) and v_1 (filled diamonds) for the 13 e.p.s.ps analysed. Data from equations (1) and (2) have been binned into 100-trial epochs, and each point shows the mean for the 13 e.p.s.ps, each normalized by its mean value during the 500-trial period before the tetanus. The change in v_1 was, on average, relatively greater than the change in P immediately after the tetanus, but it showed greater decrement over time. *b*, *c*, Graphs showing the relative enhancement of P (*b*) and v_1 (*c*) following the tetanus as a function of the initial P . The P and v_1 ratios were calculated as the mean value for the 500 trials immediately after the tetanus divided by the mean for the 500 trials before the tetanus. Each point represents an individual e.p.s.p. For *c*, least-squares linear regression gave $r^2=0.716$ ($P<0.001$). e.p.s.ps with low initial P values show greater increases in P , but smaller increases in v_1 . e.p.s.ps recorded in the lower Ca^{2+} concentration had generally low initial P values and showed only minor changes in v_1 .

release. We grouped the e.p.s.p. amplitudes into epochs of 100–250 consecutive trials and calculated the mean and variance (corrected for contaminating noise) and hence the coefficient of variation (CV) for each. First, we used these to plot $1/CV^2$ against mean, as introduced to the long-term potentiation (LTP) field by Malinow and Tsien⁸. For a wide range of models of release^{8,9} (but not all^{15,16}), trajectories as steep as or steeper than the diagonal in these plots indicate a change in the number of quanta released, whereas trajectories close to the horizontal indicate a change in postsynaptic responsiveness.

To obtain quantitative information from the mean and variance data requires the assumption of a more constrained model of the release process. We used a simple binomial model, for which

$$\text{Mean} = nPv_1 \quad (1)$$

$$\text{Variance} = nP(1-P)v_1^2 \quad (2)$$

where n is the number of sites releasing quanta, v_1 is the amplitude of the postsynaptic response to each quantum and P is the probability of any given site releasing a quantum on each trial³. Unique solutions can only be obtained by constraining one of these parameters to be constant over time. We assumed that n remained constant. To estimate n for each e.p.s.p., we first derived a value for v_1 from an amplitude histogram for a given period, then substituted this and the e.p.s.p. mean and variance for that period into equations (1) and (2) to give P and n . We then assumed n remained constant and used that value together with measured mean and variance for all the epochs of data to calculate P and v_1 throughout. The v_1 predictions from this simple model were in good agreement with quantal size estimates obtained from any additional peaky amplitude histograms that were available for each e.p.s.p.

Thirteen e.p.s.ps that showed an enhancement persisting for at least 15 min after tetanic stimulation were analysed as above. Two examples showing very different behaviour are illustrated in Fig. 3. Panels *a*–*e* show an e.p.s.p. that increased from about 90 to 210 μV (*a*). The amplitude histograms (*b*) show that the peak separation increased by less than 10%, indicating only a minor increase in quantal size. The relative heights of the peaks changed considerably, with the left-hand peak (in this case representing failures) becoming relatively lower whereas the

other peaks became higher. This indicates an increase in the mean number of quanta released per trial. The trajectory of the $1/CV^2$ plot (*c*) is steeper than the diagonal, indicating a predominantly presynaptic change. Using equations (1) and (2) with $n=3$ showed that P increased from about 0.2 to 0.4 (*d*), whereas there was only a small increase in v_1 (*e*). Thus the various analysis procedures were in good agreement, and the enhancement in this case was predominantly presynaptic. The e.p.s.p. in panels *f*–*j* behaved very differently. The amplitude increased from about 500 to 950 μV then gradually declined (*f*). The histograms (*g*) show an increase in peak separation from 135 to 230 μV . The relative heights of the histogram peaks were virtually identical in the two cases, indicating no appreciable change in the mean number of quanta released per trial. The trajectory of the $1/CV^2$ plot after the tetanus was horizontal (*h*), indicating a postsynaptic change. Using equations (1) and (2) with $n=5$ showed that v_1 (*j*) increased from 125 to 250 μV , then declined, whereas P remained almost constant (*i*). Thus these analysis procedures indicate that this e.p.s.p. showed essentially a purely postsynaptic change. Of the other e.p.s.ps analysed, some showed changes that were mainly pre- or postsynaptic, others a mixture of both. This confirms that both factors can contribute to LTP in CA1, as shown by Kullmann and Nicoll using different techniques¹⁴.

How valid are our assumptions? We are not suggesting that the good alignment we observed between the results of using equations (1) and (2) and the analysis of amplitude histograms means that release probabilities are identical at all sites or that n is always invariant. The alignment of predicted v_1 values with histogram peak spacings makes major changes in n unlikely, but minor changes cannot be ruled out. Thus we propose that the simple binomial model with constant n is a useful simplification of what is likely to be a more complex process.

What factors might determine the relative contribution of pre- and postsynaptic mechanisms at a given synapse? The mean changes in P and v_1 for our experimental set are shown in Fig. 4*a*. In Fig. 4*b*, *c* the relative P and v_1 changes are plotted as a function of the initial P for each e.p.s.p. It is clear that e.p.s.ps with a low initial P were likely to show a large change in P , whereas those with a high initial P showed larger v_1 changes. The conflicting results in the literature about the pre-

or postsynaptic locus for the maintenance of LTP may be resolved by this data. Recent studies showing predominantly presynaptic LTP have used 2.5 mM Ca^{2+} for slice experiments^{8,9,11}, and some have reported a high proportion of failures^{8,11}, suggesting low initial release probabilities. Our e.p.s.ps recorded in 2.5 mM Ca^{2+} also showed low initial P and

mainly presynaptic LTP. Other studies reporting postsynaptic LTP have used higher Ca^{2+} levels^{17,18}. Our results in higher Ca^{2+} showed higher P values and larger v_1 changes. We expect that this clear correlation between initial release probability and the locus of change will be important for the further understanding of the mechanism of LTP at these synapses. □

Received 26 May; accepted 30 September 1992.

1. del Castillo, J. & Katz, B. *J. Physiol., Lond.* **124**, 560–573 (1954).
2. Redman, S. *Physiol. Rev.* **70**, 165–198 (1990).
3. Korn, H. & Faber, D. S. *Trends Neurosci.* **14**, 439–445 (1991).
4. Larkman, A., Stratford, K. & Jack, J. *Nature* **350**, 344–347 (1991).
5. Clements, J. *Nature* **353**, 396 (1991).
6. Larkman, A., Stratford, K. & Jack, J. *Nature* **353**, 396 (1991).
7. Voronin, L. L. *Neuroscience* **10**, 1051–1069 (1983).
8. Malinow, R. & Tsien, R. W. *Nature* **346**, 177–180 (1990).
9. Bekkers, J. M. & Stevens, C. F. *Nature* **346**, 724–729 (1990).
10. Foster, T. C. & McNaughton, B. L. *Hippocampus* **1**, 79–91 (1991).
11. Malinow, R. *Science* **252**, 722–724 (1991).

12. Manabe, T., Renner, P. & Nicoll, R. A. *Nature* **355**, 50–55 (1992).
13. Malgaroli, A. & Tsien, R. W. *Nature* **357**, 134–139 (1992).
14. Kullmann, D. M. & Nicoll, R. A. *Nature* **357**, 240–244 (1992).
15. Korn, H., Fassnacht, C. & Faber, D. S. *Nature* **350**, 282 (1991).
16. Faber, D. S. & Korn, H. *Biophys. J.* **60**, 1288–1294 (1991).
17. Muller, D., Joly, M. & Lynch, G. *Science* **242**, 1694–1697 (1988).
18. Kauer, J. A., Malenka, R. C. & Nicoll, R. A. *Neuron* **1**, 911–917 (1988).
19. Magleby, K. L. & Miller, D. C. *J. Physiol., Lond.* **311**, 267–287 (1981).
20. Edwards, F. A., Konnerth, A. & Sakmann, B. *J. Physiol., Lond.* **430**, 213–249 (1990).

ACKNOWLEDGEMENTS. This work was supported by the Wellcome Trust. T.H. is a Wellcome Prize Student and A.L. is a Royal Society University Research Fellow.

Preserved figure-ground segregation and symmetry perception in visual neglect

Jon Driver, Gordon C. Baylis* & Robert D. Rafal†

Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

* Department of Psychology 0109, University of California, San Diego, La Jolla, California 92093, USA

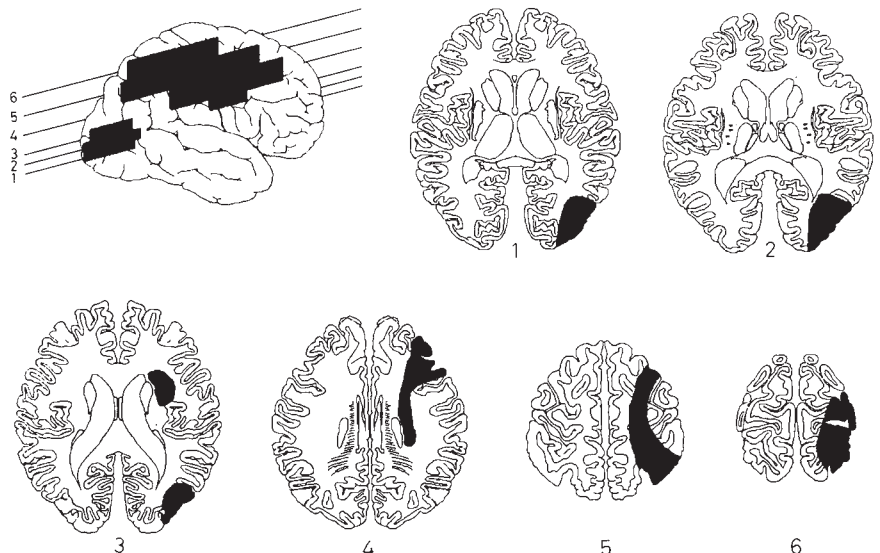
† Department of Neurology and Center for Neurobiology, University of California, Davis, California 94553, USA

A CENTRAL controversy in current research on visual attention is whether figures are segregated from their background preattentively^{1–8}, or whether attention is first directed to unstructured regions of the image^{9–13}. Here we present neurological evidence for the former view from studies of a brain-injured patient with visual neglect. His attentional impairment arises after normal segmentation of the image into figures and background has taken place. Our results indicate that information which is neglected and unavailable to higher levels of visual processing can nevertheless be processed by earlier stages in the visual system concerned with segmentation.

C.C. is a 69-year-old man with right-hemisphere damage and severe left neglect^{14,15} (see Fig. 1 for case details). He ignores information on the left even though he has no left visual field loss. His neglect can therefore be characterized as an attentional rather than sensory deficit. We examined whether this deficit applied to the left side of unsegregated space, as would be predicted by purely spatial models of visual attention^{9–13}, or to the left of perceptual figures derived by preattentive segregation processes, as would be predicted by object-based models of attention^{1–8}. He was presented with displays such as those in Fig. 2. A rectangle presented on a black screen was divided by a pseudorandom contour into a small bright green section and a large dimmer red section. The green section could be on the far left (upper panel of Fig. 2) or the far right (lower panel of Fig. 2). With appropriate colouring, normal observers see such displays as green shapes against a shapeless red background. This figure-ground segregation arises because the green shapes are smaller and brighter¹⁶.

C.C.'s task was to remember the dividing contour between the red and green section on each trial, deciding whether it matched a probe line presented after each display (for example, in Fig. 2 the probe matches for the upper panel but not the lower panel). If his attentional deficit applies to the left side of an unsegregated representation of visual space, C.C. should perform less well when the green shape appears at the left of the rectangle (as in the upper panel of Fig. 2) as the dividing

FIG. 1 Reconstructed computerized tomographic scan for C.C., a right-handed retired plumber who suffered a right-hemisphere stroke and sustained left hemiparesis and left neglect. The method for reconstruction is described elsewhere²⁶. A right lateral view and transverse sections are shown, with shading indicating damage. The scan reveals infarction of the watershed territory of the right anterior and middle cerebral arteries. The damage extends anteriorly into area 8 and posteriorly into the superior parietal lobule. C.C. had no visual-field loss, but consistent left-sided extinction on double simultaneous stimulation. He showed severe left neglect on the conventional measures of line bisection and line cancellation, and in a new measure; when asked to detect the presence or absence of a small crack on either side of two-dimensional block shapes, he detected the crack for 37/40 shapes with a crack on the right, but for only 5/40 shapes with a crack on the left. All studies were conducted in the month following the stroke in the sequence described.



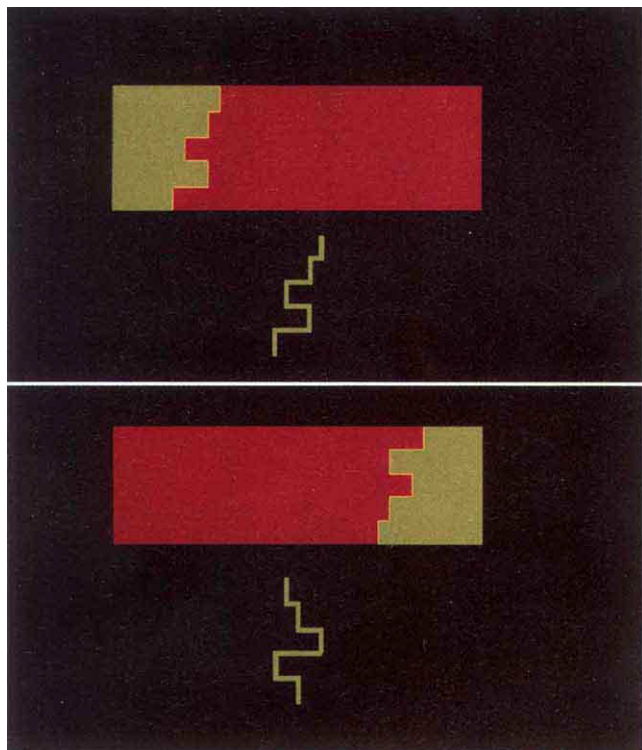


FIG. 2 Two example displays from the first experiment. The task was to report verbally whether the dividing contour in the $9.5^\circ \times 3.2^\circ$ rectangle matched the probe line, which was presented immediately below the rectangle following its offset as shown. The smaller region of the rectangle (on average 3.2° wide) was always bright green (~ 3.8 footlamberts (ftL)), the larger region a dimmer red (2.4 ftL) against a black background (0.15 ftL). In the upper panel, the small green figure is on the left, and a matching probe is illustrated. In the lower panel the small green figure is on the right, and a mismatching probe is illustrated. The green figure was equally likely to be on the left or right, and 50% of the probes matched in both cases. The dividing contours and probes always had 5 pseudorandom 'steps'. Using an IBM PC-AT microcomputer and a Mini-Micro VGA monitor, the rectangle was presented for 2 s, followed by a gap of 100 ms, and then the probe until a same/different response was made. Central fixation was required during each rectangle display. There were 160 trials in random sequence. C. C. correctly rejected mismatching probes on 33/40 trials when the green figure was on the left, but on only 21/40 trials with green figures on the right. He correctly accepted matching probes on 31/40 trials when the green figure was on the left, and on 19/40 trials with green figures on the right. Thus, C.C. performed well above chance for displays with the green figure on the left, as in the upper panel, but was at chance when the green figure was on the right, as in the lower panel. This difference in performance is significant ($P < 0.01$).

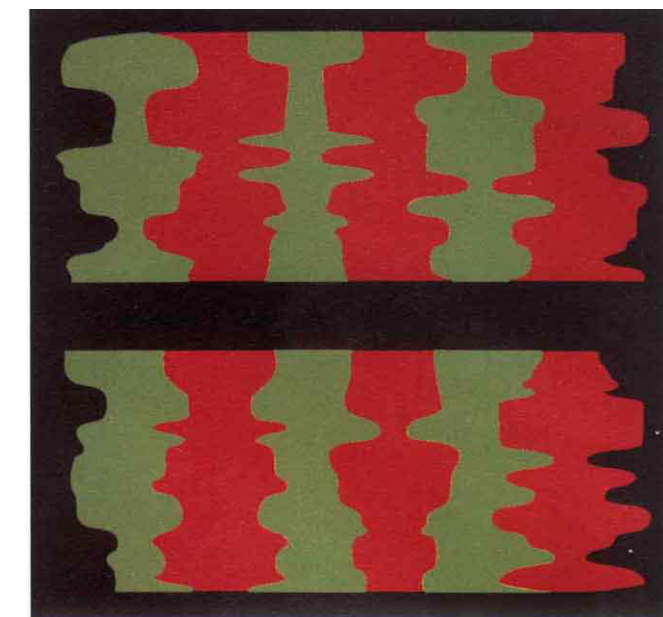
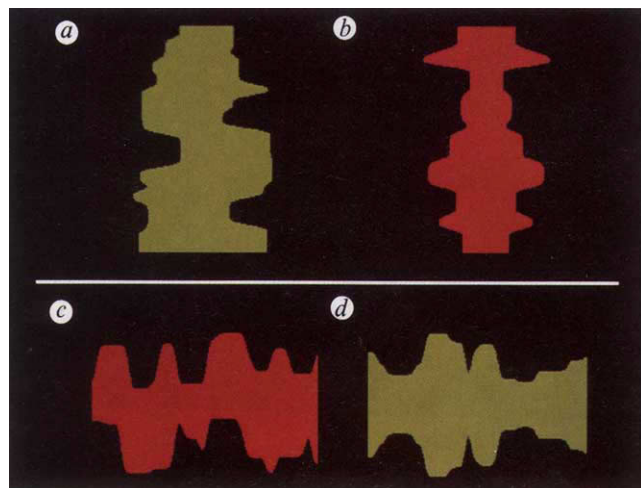


FIG. 3 Two examples of displays from the second study. Readers should be able to experience the effect of symmetry on figure-ground segregation for themselves. With the lower example covered, the upper figure is typically seen as green shapes against a red background because of the symmetry. Conversely, with the upper example covered, the lower is typically seen as red figures against a green ground. Each display comprised 6 pseudorandom shapes, alternating red and green, and subtended $13.4^\circ \times 6.3^\circ$ in total. With equal probability, either the red shapes were three different symmetrical shapes, and the green shapes asymmetrical, or the reverse. The rightmost shape was equally likely to be symmetrical or asymmetrical, and to be red (4.3 ftL) or green (3.8 ftL). The colours were roughly matched for salience in pilot studies with healthy observers. The displays were presented until a response was made, and there were 120 trials. With red symmetrical shapes, C.C. responded 'red' on 21/30 trials when the rightmost shape was red, and 26/30 trials when the rightmost shape was green. With green symmetrical shapes, he responded 'green' on 29/30 trials when the rightmost shape was green, and 20/30 trials when the rightmost shape was red. Thus, there was no significant tendency to pick one colour rather than another (58/120 red responses, 62/120 green responses), or to respond with the colour of the rightmost shape (64/120), but a strong tendency to pick the colour of the symmetrical shapes (96/120, $P < 0.01$). This effect of symmetry was indistinguishable from that observed in the speeded responses of two age-matched, neurologically healthy men. The colour of the asymmetrical shapes is chosen on a small proportion of trials where other factors (such as relative size) override symmetry because of pseudorandom fluctuation in the shapes employed.

FIG. 4 *a, b*, Two example displays from the explicit vertical symmetry task. Each display comprised a single shape from the figure-ground displays of the preceding study. The shapes were 6.3° tall and on average 2.2° wide, and were equally likely to be red or green. In either case, 50% were symmetrical. A green asymmetrical shape (*a*) and a red symmetrical shape (*b*) are shown. Individual displays were presented until a response was made, and there were 120 trials. C.C.'s performance was at chance; he responded 'symmetrical' to 31/60 symmetrical shapes, and 'asymmetrical' to 34/60 asymmetrical shapes, with no effect of display colour. *c, d*, Two examples of displays from the explicit horizontal symmetry task, conducted immediately after C.C.'s chance performance in judging vertical symmetry. A red asymmetrical shape (*c*) and a green symmetrical shape (*d*) are shown. The displays were exactly as for the vertical symmetry task, except that the monitor was rotated 90° clockwise. Normal observers find symmetry judgements harder around the horizontal than vertical²⁷. But C.C.'s abnormal difficulty in judging vertical symmetry is caused by his neglect for the left of figures, which should not disrupt judgements of symmetry about the horizontal axis. C.C. responded 'symmetrical' to 56/60 symmetrical shapes, and 'asymmetrical' to 59/60 asymmetrical shapes. This performance is significantly better than that for vertical symmetry ($P < 0.01$).

contour is further to the left of the display and the patient in this case. But the opposite prediction follows if C.C.'s neglect applies to the left side of perceptual figures after the normal operation of figure-ground segregation. He should perform worse when the green shape is on the right of the rectangle (as in the lower panel of Fig. 2) because the dividing contour now falls to the left of the perceptual figure, although the contour is further to the right of the display and the patient.

The results supported the figural prediction (see legend to Fig. 2), demonstrating that the critical variable was the location of the dividing contour in relation to the green figure, rather than in relation to the patient or to the display as a whole. This implies that, like normal observers, C.C. saw the green shapes as the perceptual figures whether they appeared on the right or the left, implying that figure-ground segregation operated normally across the display (his phenomenal descriptions of 'green shapes at either side' on debriefing are also consistent with this). Thus, C.C.'s left neglect arises at a stage of attending to figures provided by a preceding, and unimpaired, segregation stage. This two-stage account of his neglect is in accordance with recent two-stage characterizations of normal vision⁸. Our results are consistent with a recent report that visual neglect can apply to the contralateral side of objects¹⁷. They go beyond this by demonstrating that the 'objects' concerned are figures derived by segregation processes operating normally in both visual fields. The intact segregation between green and red shapes is also consistent with a previous report that colour perception can be preserved in the contralateral field for some neglect patients¹⁸. Our data extend this to show that such preserved perception can support figure-ground segregation.

Our next experiment tested directly whether C.C.'s visual system represents left-sided information at a preattentive stage of figure-ground segregation, but subsequently loses this information for the derived figures at an attentional stage. Symmetry about the vertical exerts powerful effects on figure-ground segregation in normal observers; other factors being equal, symmetrical shapes are seen as figures against asymmetrical grounds¹⁶. We examined whether C.C. also showed this effect. Because symmetry depends on the correspondence between reflected sides, sensitivity to symmetry requires representation of both the right and left sides of shapes at the segregation stage.

We presented C.C. with displays of alternating red and green pseudorandom shapes (Fig. 3). Either the red shapes or the green shapes were symmetrical. C.C. was told: "Sometimes we are going to show you red shapes on a green background, sometimes green shapes on a red background. Just tell us which colour the shapes are, red or green". He showed the usual effect of symmetry on figure-ground segregation, tending to report

the colour of the symmetrical shapes, with no bias towards reporting the rightmost colour (see legend to Fig. 3).

This normal effect of symmetry demonstrates that, at the preattentive segregation stage, both sides of perceptual figures must be represented in C.C.'s visual system. But our first experiment found that he neglects the left side of perceptual figures at a subsequent stage of attending to them. C.C.'s phenomenal descriptions of the present displays suggest that this applied for the symmetrical figures as well. When asked to explain his responses at the end of the study, he replied that his chosen shapes 'looked closer and brighter', but did not mention the symmetry which is so striking to a normal observer. In the next task, we required C.C. to judge explicitly the symmetry of individual shapes from the figure-ground displays (Fig. 4a and b). His responses were at chance level (see legend to Fig. 4). This inability to make explicit symmetry judgements contrasts dramatically with the preserved effect of symmetry on figure-ground segregation. The contrast arises because the two tasks reflect distinct stages of visual processing. Figure-ground segregation occurs preattentively, at a stage where left-sided information is represented. Explicit symmetry judgements require a subsequent stage of attending to perceptual figures, at which the patient neglects left-sided information. His inability to make explicit vertical symmetry judgements was not due to failure to understand the task, as he was able to judge whether our non-sense shapes were symmetrical about the horizontal axis (Fig. 4c and d). This task requires comprehension of comparable instructions, but does not require left-sided information.

C.C. retains intact figure-ground segregation despite his severe left neglect. He segregates figures from their background normally, even when they appear on his left (our first study), and even when their segregation depends on vertical symmetry (our second study) which requires that both the right and left sides of these figures are represented. His deficit arises in a subsequent stage of attending to these figures, at which he neglects the left of figures wherever they appear, and thus cannot judge explicitly whether they are symmetrical about the vertical.

The dissociation between severe left neglect and intact figure-ground segregation strongly supports recent models of normal visual function, which postulate a preattentive segregation stage providing candidate objects to a subsequent attentional stage¹⁻⁸. Our findings about symmetry perception also demonstrate that neglected information that is unavailable for verbal report may nevertheless be represented implicitly, confirming previous reports^{19,20} which have remained controversial²¹⁻²³. This conclusion is in accordance with comparable dissociations between implicit processing and verbal report in other areas of neuropsychology^{24,25}. □

Received 19 June; accepted 28 September 1992.

1. Baylis, G. C. & Driver, J. *Percept. Psychophys.* **51**, 145-162 (1992).
2. Driver, J. & Baylis, G. C. *J. exp. Psychol. hum. Percept. Perform.* **15**, 448-456 (1989).
3. Duncan, J. *J. exp. Psychol. Gen.* **13**, 501-517 (1984).
4. Julesz, B. *Trends Neurosci.* **6**, 41-45 (1984).
5. Kahneman, D., Treisman, A. & Gibbs, B. J. *Cognitive Psychol.* **24**, 175-219 (1992).
6. Kanwisher, N. & Driver, J. *Curr. Dir. psychol. Sci.* **1**, 26-31 (1992).
7. McLeod, P., Driver, J. & Crisp, J. *Nature*, **332**, 154-155 (1988).
8. Treisman, A. Q. J. *exp. Psychol.* **40A**, 201-237 (1988).
9. Crick, F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4586-4590 (1984).
10. Posner, M. I. *J. exp. Psychol.* **32**, 3-25 (1980).
11. Tsai, Y. & Lavie, N. *Percept. Psychophys.* **44**, 15-21 (1988).
12. Umiltà, C. in *Handbook of Neuropsychology* Vol. 1 (eds Boller, F. & Grafman, J.) 175-193 (Elsevier, Amsterdam, 1989).
13. van der Heijden, A. H. C. *Selective Attention in Vision* (Routledge, London, 1992).
14. Heilman, K. M. et al. in *Clinical Neuropsychology* (eds Heilman, K. M. & Valenstein, E.) 243-293 (Oxford Univ. Press, New York, 1985).

15. Bisiach, E. & Vallar, G. in *Handbook of Neuropsychology* Vol. 1 (eds Boller, F. & Grafman, J.) 195-122 (Elsevier, Amsterdam, 1989).
16. Rock, I. *The Logic of Perception* (MIT Press, Cambridge, Massachusetts, 1983).
17. Driver, J. & Halligan, P. W. *Cognitive Neuropsychol.* **8**, 475-496 (1991).
18. Cohen, A. & Rafal, R. D. *Psychol. Sci.* **2**, 106-110 (1991).
19. Volpe, B. T., Ledoux, J. E. & Gazzaniga, M. S. *Nature* **282**, 722-724 (1979).
20. Sieroff, E. et al. *Cognitive Neuropsychol.* **5**, 427-429 (1988).
21. Farah, M. J. et al. *Neuropsychologia* **29**, 949-958 (1991).
22. Brunn, J. L. & Farah, M. J. *Cognitive Neuropsychol.* **8**, 59-75 (1991).
23. Berti, A. et al. *Neuropsychologia*, **30**, 403-415 (1992).
24. Rafal, R. D. et al. *Science* **250**, 118-121 (1990).
25. Tranel, E. & Damasio, A. R. *Science* **228**, 1453-1454 (1985).
26. Robertson, L. C. et al. *J. Neurosci.* **8**, 3757-3769 (1988).
27. Barlow, H. B. & Reeves, B. C. *Vision Res.* **19**, 783-793 (1979).

ACKNOWLEDGEMENTS. We thank C.C. for his willing participation. The McDonnell-Pew Foundation, US Office of Naval Research, Medical Research Council (UK) and NATO provided support.

Calcium-dependent immediate feedback control of inositol 1,4,5-trisphosphate-induced Ca^{2+} release

Masamitsu Iino & Makoto Endo

Department of Pharmacology, Faculty of Medicine, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

THE temporal and spatial distribution of increases in intracellular Ca^{2+} concentration is an important factor in cellular signal transduction¹. Inositol 1,4,5-trisphosphate (InsP_3) plays a key part in agonist-induced Ca^{2+} release¹, which can take place abruptly² and in a confined space³ by a mechanism that is not fully understood. Here we analyse the kinetics of InsP_3 -induced Ca^{2+} release following flash photolysis⁴⁻⁷ of caged InsP_3 or caged Ca^{2+} , and demonstrate that Ca^{2+} -dependent immediate feedback control is an important determinant of the time course of Ca^{2+} release. The positive feedback mechanism is also important for the 'loading dependence' of InsP_3 -induced Ca^{2+} release⁸. Furthermore, our results support the operation of positive cooperativity in channel opening⁹ and feedback control augments the steep InsP_3 concentration- Ca^{2+} release relation. These inherent properties of InsP_3 -induced Ca^{2+} release are expected to give rise to temporally abrupt and/or spatially confined Ca^{2+} release within the cell.

Fluorimetry of fluo-3 (ref. 10) was used to measure Ca^{2+} release from permeabilized thin smooth muscle¹¹. Photolysis of caged InsP_3 resulted in a rapid release of Ca^{2+} in the absence of Mg-ATP^{2+} , when there is no Ca^{2+} reuptake (Fig. 1a). The overall time course was S-shaped: lag phase followed by fast and then slow release phases. The maximum extent of Ca^{2+} release was similar to that obtained by the application of 30 μM InsP_3 (not shown), which releases ~80% of ionomycin-releasable Ca^{2+} in this preparation¹¹. The time constant for the liberation of InsP_3 (~10 ms)⁴ is shorter than the delay. Neither flash without caged InsP_3 nor photolysis of caged ATP induced Ca^{2+} release.

The steady state rate of InsP_3 -induced Ca^{2+} release depends biphasically on the Ca^{2+} concentration, with the peak near 300 nM^{12,13}. Because Ca^{2+} release was initiated at 110 nM Ca^{2+} (Fig. 1a), almost the resting intracellular concentration, the Ca^{2+} dependence of InsP_3 -induced Ca^{2+} release was expected to give rise to positive feedback in the early phase of Ca^{2+} release. The late slowing of the Ca^{2+} release could have been due to higher Ca^{2+} -dependent inhibition. Although the feedback control seems to be consistent with the observed time course, we did not know whether the feedback was rapid and efficient enough to explain the observations. We therefore designed experiments to test whether the feedback mechanisms control the kinetics of Ca^{2+} release.

A higher concentration of Ca^{2+} buffer is expected to make the positive-feedback loop less effective. The time course indeed became slower and less sigmoidal with 200 μM fluo-3 (Fig. 1b). In this preparation, a Ca^{2+} -induced Ca^{2+} release (CICR) mechanism is also present in ~20% of InsP_3 -sensitive Ca^{2+} stores¹¹. But, essentially the same sigmoidal time course was obtained (Fig. 1c) in the presence of 50 μM ruthenium red, an inhibitor of CICR¹⁴. Thus, CICR is not essential for the S-shaped time course.

When caged InsP_3 was photolysed at a higher pre-flash Ca^{2+} concentration ($[\text{Ca}^{2+}]_{\text{init}}$), that is, near the peak of the biphasic Ca^{2+} dependence of InsP_3 -induced Ca^{2+} release, the lag time became much shorter, the initial rate of Ca^{2+} release greater, and the slowing of Ca^{2+} release started earlier (Fig. 2a). These results suggest that feedback control of Ca^{2+} release is the underlying basis of the sigmoidal time course. Similar results

were obtained in the presence of 0.5 mM Mg^{2+} (not shown).

If there is feedback control of InsP_3 -induced Ca^{2+} release, then the rate of Ca^{2+} release may depend on the extent of Ca^{2+} loading of the Ca^{2+} store. If the Ca^{2+} store is only partially loaded with Ca^{2+} , the feedback control is likely to be less effective. The time course of InsP_3 -induced Ca^{2+} release was, indeed, slower with lower Ca^{2+} loading, as can be clearly seen after scaling of the signal to match the peak-release sizes (Fig. 2b). Even with low Ca^{2+} loading there was clear dependence of the rate of Ca^{2+} release on $[\text{Ca}^{2+}]_{\text{init}}$ (Fig. 2c). Hence the positive feedback is an important basis of 'loading dependence' of InsP_3 -induced Ca^{2+} release⁸. It could be that the intravesicular Ca^{2+} concentration is also controlling the Ca^{2+} release rate¹⁵, but the luminal Ca^{2+} dependence alone cannot entirely explain our results.

To see how rapidly the effect of Ca^{2+} on InsP_3 -induced Ca^{2+} release appears, we studied the effect of a Ca^{2+} -concentration jump created by photolysis of DM-nitrophen⁷. If the Ca^{2+} concentration is suddenly increased along the ascending part of the biphasic dependence of InsP_3 -induced Ca^{2+} release (from ~50 nM to ~165 nM), there is an immediate potentiation of the rate of Ca^{2+} release (Fig. 3a). In the case of a Ca^{2+} jump on the descending part (from ~270 nM to ~850 nM), there is an immediate slowing of the rate of Ca^{2+} release (Fig. 3b). Thus Ca^{2+} immediately controls the rate of InsP_3 -induced Ca^{2+} release. The concentration of free Ca^{2+} in the immediate vicinity of the Ca^{2+} stores would have been somewhat greater than the average $[\text{Ca}^{2+}]_{\text{free}}$ (Fig. 3b) at the time of photolysis because there was a significant Ca^{2+} release, but this should not affect the conclusion. Although Ca^{2+} -dependent inhibition of InsP_3 -induced Ca^{2+} release has been reported to develop with a time constant of 0.58 s in synaptosomes¹⁶, a possible reduction of intravesicular Ca^{2+} through CICR during the preincubation might explain the apparently slow kinetics.

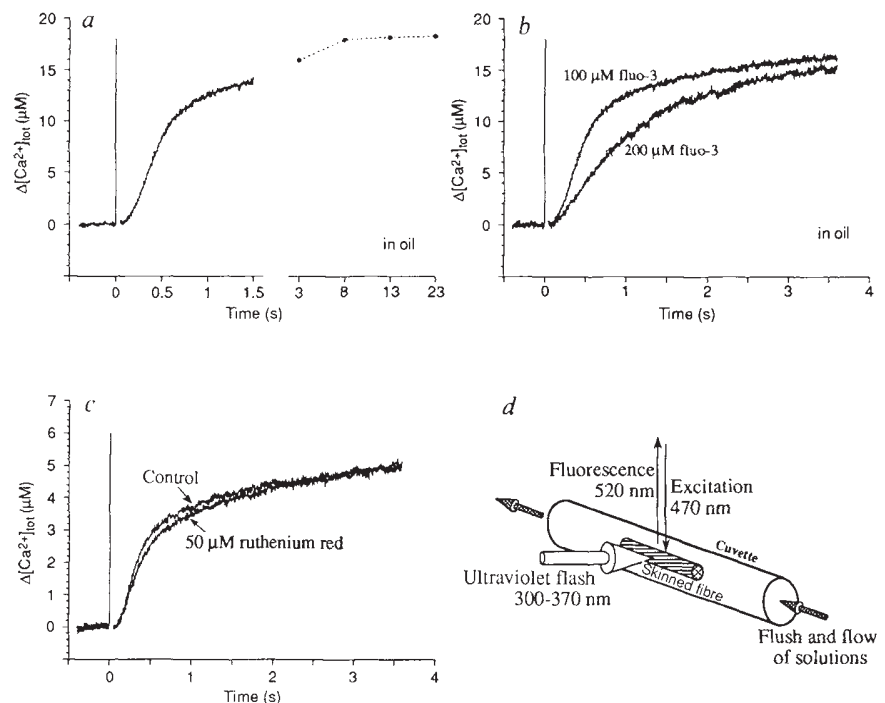
We examined the effects of feedback control on the dose-response relation of InsP_3 -induced Ca^{2+} release. As shown in Fig. 4a, the Ca^{2+} release depends very steeply on InsP_3 concentration when feedback control is allowed. In Fig. 4c (filled circles) the rates of Ca^{2+} release are compared at the time when the Ca^{2+} release rate reaches the maximum with the highest InsP_3 concentration (downward arrow in Fig. 4a). The maximum rates of Ca^{2+} release at lower flash intensities were obtained later (arrowheads in Fig. 4a; triangles in Fig. 4c). The steep dose-response relation of InsP_3 -induced Ca^{2+} release may result from positive-feedback control or cooperative channel opening, or both. We therefore studied the InsP_3 -concentration dependence at a higher $[\text{Ca}^{2+}]_{\text{init}}$ in the presence of a higher fluo-3 concentration and with lower Ca^{2+} loading to exclude the feedback control. The results (Fig. 4b) are inevitably noisier but clearly show a monotonic time course. The initial rate of Ca^{2+} release still had cooperative dependence on the InsP_3 concentration (Fig. 4c, circles; Hill coefficient ~2). Therefore positive cooperativity may be operational and positive feedback control augments the steepness of the dose-response relation.

Positive cooperativity of InsP_3 -induced Ca^{2+} release with a Hill coefficient of 3 to 4 has been reported in permeabilized basophilic leukaemia cells⁹. In the earlier experiments using a weak Ca^{2+} buffer (1.5 μM Fura-2), positive cooperativity may have been overestimated because of positive feedback control but it could still have been a real difference. In contrast, InsP_3 concentration dependence of the open probability of InsP_3 -gated channels incorporated into planar lipid bilayers shows no cooperativity¹⁷. The study did not discriminate the subconductance levels when calculating the open probability, and it is possible that distribution of the subconductance level is dependent on the InsP_3 concentration and the time-averaged conductance may have a steeper InsP_3 dependence.

Owing to the positive feedback control and positive cooperativity of the InsP_3 -induced Ca^{2+} release, there is a narrow range of InsP_3 concentration that can trigger an abrupt release of Ca^{2+} .

FIG. 1 *a*, Time course of Ca^{2+} release from a permeabilized smooth muscle preparation following flash photolysis (at zero time) of $1\ \mu\text{M}$ caged InsP_3 in the presence of $100\ \mu\text{M}$ fluo-3. Note the double time base. *b*, Comparison of Ca^{2+} release with $100\ \mu\text{M}$ (same experiment as in *a*) and $200\ \mu\text{M}$ fluo-3. $[\text{Ca}^{2+}]_{\text{init}}$ was $110\ \text{nM}$ in both cases. Records are representative of 5 experiments. *c*, Ca^{2+} release with or without $50\ \mu\text{M}$ ruthenium red with $50\ \mu\text{M}$ fluo-3 (3 experiments). *d*, Schematic drawing of the experimental set-up, see Methods.

METHODS. Ca^{2+} release was measured fluorometrically^{11,12,23}. Saponin-permeabilized smooth muscle strips (thickness $\sim 70\ \mu\text{m}$, width $\sim 200\ \mu\text{m}$, length $2.5\ \text{mm}$) from guinea-pig portal vein were attached to a stainless steel wire (diameter $100\ \mu\text{m}$) and mounted in a capillary cuvette (internal diameter $400\ \mu\text{m}$) through which solutions were perfused (*d*). The full width of the cuvette at the centre of the specimen was viewed by a long working distance objective (Olympus, ELWD $20\times$, NA 0.4) and the output of the fluorometer (CAM-200, JASCO, Japan) attached to an epifluorescence microscope was low-pass filtered (6-pole Chebyshev, $f_c = 200\ \text{Hz}$) before sampling at $1\ \text{kHz}$. Defocusing of the microscope by $200\ \mu\text{m}$ resulted in only a 5% change in fluorescence intensity, so the fluorometer detected Ca^{2+} change throughout the cuvette. Light emitted from a xenon flash lamp (SA-200E; Eagle, Japan) was filtered (300–370 nm) for photolysis and had a decay constant of $\sim 0.8\ \text{ms}$, but the ringing of the low-pass filter prevented measurement for $\sim 40\ \text{ms}$ after the flash. Decay of the intensity of the photolysis light along the light pass was negligible except for the backside of the stainless steel wire. The specimen was placed on the front-side of the wire for both photolysis and fluorometry. There was linear relation between total added Ca^{2+} and the fluorescence intensity of fluo-3 up to about $2/3$ of the maximum change (compare Fig. 1 of ref. 23) with or without the preparations being present. The fluorescence intensity change of fluo-3 scaled by the maximum change and multiplied by the dye concentration was therefore taken as $\Delta[\text{Ca}^{2+}]_{\text{tot}}$, which should be proportional to the total amount of Ca^{2+} released from the stores averaged in the cuvette or in the fibre volume (in oil, see below), although the absolute value is underestimated. The dissociation constant of fluo-3 for Ca^{2+} ($580\ \text{nM}$, determined in the cuvette) was used to estimate the average free



Ca^{2+} concentrations. This value should be the same as that in the vicinity of the Ca^{2+} stores in the steady state, but the latter may become greater than the average during rapid Ca^{2+} release. InsP_3 liberated in the cuvette was measured by an InsP_3 -binding protein kit (Amersham) and depended proportionally on the flash intensity. A $200\ \text{J}$ -flash photocleaved 10% of caged InsP_3 . After Ca^{2+} loading at $1\ \mu\text{M}$ Ca^{2+} in the presence of $4\ \text{mM}$ Mg-ATP^{2-} for 30 to 180 s, Ca^{2+} and Mg^{2+} were washed out with a solution containing $1\ \text{mM}$ EDTA (solutions 2 and 4 of ref. 11). Then fluo-3 (50 – $200\ \mu\text{M}$) and caged InsP_3 ($1\ \mu\text{M}$; Calbiochem) were introduced 60 s before the flash in a solution containing (in mM) Mg^{2+} (0); ATP (2); potassium methanesulphonate (130); NaN_3 (20); dithiothreitol (5); PIPES (20), pH 7.0 at 20 – 22°C . Mg^{2+} -free solution was used to inhibit Ca^{2+} uptake and InsP_3 phosphatase²⁴. In some experiments (indicated by 'in oil') the solution surrounding the preparation was replaced by silicone oil 20 s before the flash. This minimized the inhomogeneity in photolysis and the fluorescence measurement. Results were essentially the same except for the compression of the sigmoidal time course along the time base.

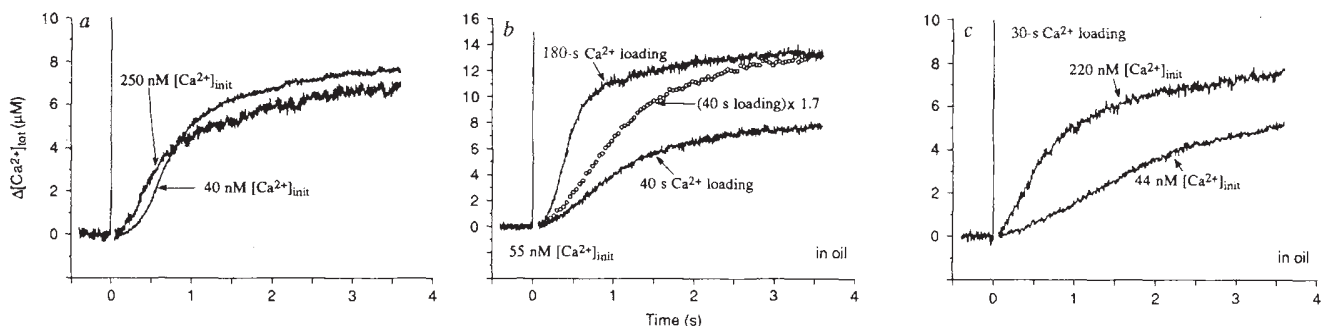
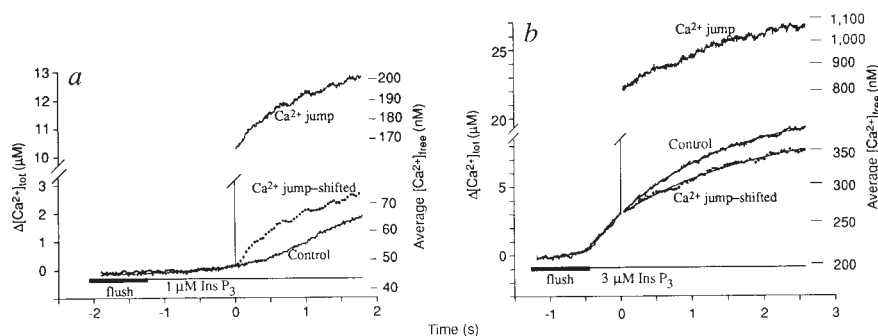


FIG. 2 Dependence of Ca^{2+} release on the pre-flash Ca^{2+} concentration (*a*, *c*) and on the amount of Ca^{2+} loaded in the stores (*b*). *a*, Time courses of Ca^{2+} release were compared with or without addition of Ca^{2+} in the presence of $100\ \mu\text{M}$ fluo-3 to make $[\text{Ca}^{2+}]_{\text{init}}$ 40 or $250\ \text{nM}$. The noise level was proportional to the fluorescence intensity. Therefore, the fluorescence intensity signal of fluo-3 at higher Ca^{2+} concentrations was always noisier

than that at lower Ca^{2+} concentrations (5 experiments). *b*, Amount of Ca^{2+} loaded in the store was varied by changing the loading time (40 s or 180 s). Circles show scaled-up data of the low Ca^{2+} loading to match the final Ca^{2+} release size; $100\ \mu\text{M}$ fluo-3, $\sim 55\ \text{nM}$ $[\text{Ca}^{2+}]_{\text{init}}$ in both conditions (9 experiments). *c*, Dependence of the kinetics of Ca^{2+} release on $[\text{Ca}^{2+}]_{\text{init}}$ with partial Ca^{2+} loading (30-s loading); $100\ \mu\text{M}$ fluo-3 (4 experiments).

FIG. 3 Effect of Ca^{2+} -concentration jump on the rate of InsP_3 -induced Ca^{2+} release. Comparison was facilitated by shifting the trace after the Ca^{2+} jump. The ordinates are broken for clarity. *a* InsP_3 -induced Ca^{2+} release was initiated with perfusion of $1\ \mu\text{M}$ InsP_3 (for 0.85 s) in the presence of $70\ \mu\text{M}$ fluo-3, $25\ \mu\text{M}$ DM-nitrophen (1-(2-nitro-4, 5-dimethoxyphenyl)*N,N,N',N'*-tetrakis(oxy-carbonyl)methyl)-1,2-ethanediamine), and $17\ \mu\text{M}$ Ca^{2+} in addition to Ca^{2+} contamination. In the continued presence of InsP_3 , the flash light (200 J-input) photolysed $\sim 40\%$ of DM-nitrophen and the average free- Ca^{2+} concentration rose from $\sim 50\ \text{nM}$ to $\sim 165\ \text{nM}$, whereupon there was an immediate increase in Ca^{2+} release rate. *b*, Following perfusion of $3\ \mu\text{M}$ InsP_3 with $50\ \mu\text{M}$ DM-nitrophen and $50\ \mu\text{M}$ added Ca^{2+} , the Ca^{2+} concentration jumped from $\sim 270\ \text{nM}$ to $\sim 850\ \text{nM}$. There was an immediate slowing of the rate of Ca^{2+} release. A single exponential was fitted to the records after the flash, and the time derivative of the fitted exponential at the flash with the Ca^{2+} jump was 64% of that of the control. Records are representative of 3 experiments in both *a* and *b*.

METHODS. DM-nitrophen (Calbiochem) was applied 60 s before the flash in a solution containing (in mM): potassium methanesulphonate (142); NaN_3



(20); dithiothreitol (5); PIPES (20), pH 7.0. The Ca^{2+} jump without InsP_3 showed slow sagging after the step increase in Ca^{2+} . This must have been due to binding of Ca^{2+} to the preparation because it was absent without the fibre. This effect was corrected for in *a* and *b* by subtraction. The dissociation constant of DM-nitrophen for Ca^{2+} increases $\sim 10^6$ -fold from $5\ \text{nM}$ upon photolysis⁴. Therefore the caged Ca^{2+} is almost fully Ca^{2+} -bound and is fully unbound before and after the photolysis, respectively, under the present experimental conditions. InsP_3 concentrations were higher here than in caged InsP_3 experiments to overcome diffusional delay.

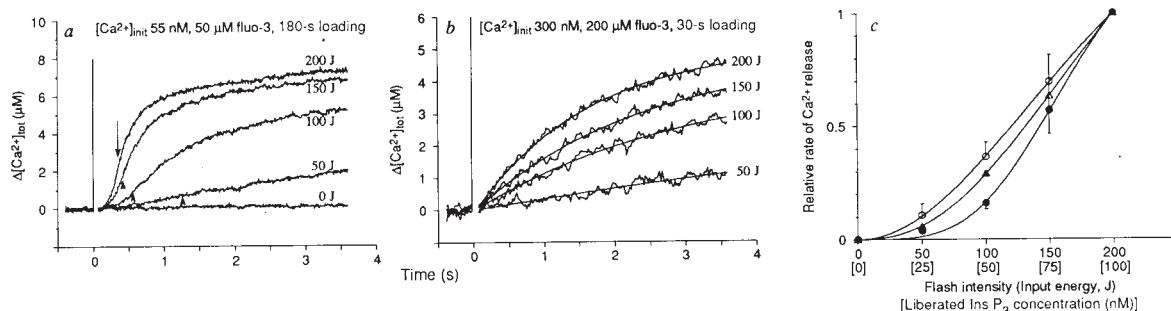


FIG. 4 Dependence of Ca^{2+} release on the concentration of InsP_3 liberated from caged InsP_3 . *a*, Flash intensity was varied by changing the energy input into the flash lamp from 50 to 200 J. The Ca^{2+} store was loaded in the presence of $1\ \mu\text{M}$ Ca^{2+} for 3 min before each run; $50\ \mu\text{M}$ fluo-3, $55\ \text{nM}$ $[\text{Ca}^{2+}]_{\text{init}}$. *b*, Experiment designed to minimize positive-feedback control. Ca^{2+} loading was for 30 s, and Ca^{2+} release was observed with $200\ \mu\text{M}$ fluo-3, $300\ \text{nM}$ $[\text{Ca}^{2+}]_{\text{init}}$. A single exponential was fitted to each record by a nonlinear least-square method. *c*, Normalized rates of Ca^{2+} release (average $\pm$ s.d.) as a function of the flash intensity. Estimated concentrations of InsP_3 liberated are indicated in brackets. Filled circles, rate of Ca^{2+} release

at the time when maximum Ca^{2+} release rate was obtained with a 200-J flash (downward arrow in *a*; $n=3$). Triangles, maximum Ca^{2+} release rate at each flash intensity (arrowheads in *a*; $n=3$). Circles: time derivative of the fitted exponential at the time of flash in same experiments as in *b*; $n=5$. The following equation was fitted to the data points using a nonlinear least-square fitting program: $Y = Y_{\text{max}} / (1 + (K/X)^h)$, where X , Y , K and h represent released InsP_3 concentration, rate of Ca^{2+} release, apparent K_M and Hill coefficient, respectively. Best-fit parameters were $K=112\ \text{nM}$, $h=2.03$ (circles); $K=113\ \text{nM}$, $h=2.41$ (triangles); and $K=88\ \text{nM}$, $h=3.82$ (filled circles).

The steep InsP_3 dependence may be useful for signal amplification. If there is a gradient of InsP_3 concentration within the cell, only the part where InsP_3 concentration exceeds a critical level may respond by releasing Ca^{2+} , as shown for the *Xenopus* oocyte⁴. Even with a homogeneous InsP_3 con-

centration, a focal rise in Ca^{2+} concentration will trigger Ca^{2+} release. This may propagate to form a Ca^{2+} wave¹⁸⁻²². The InsP_3 receptor/channels must therefore be able inherently to initiate a rapid rise in Ca^{2+} concentration or to induce spatially organized Ca^{2+} release. □

Received 27 May; accepted 11 September 1992.

- Berridge, M. J. & Irvine, R. F. *Nature* **341**, 197-205 (1989).
- Miyazaki, S. *J. Cell Biol.* **106**, 345-353 (1988).
- Parker, I. & Ivorra, I. *Science* **250**, 977-979 (1990).
- McCray, J. A. & Trentham, D. R. *Rev. Biophys. biophys. Chem.* **18**, 239-270 (1989).
- Somlyo, A. P. & Somlyo, A. V. *Rev. Physiol.* **52**, 857-874 (1990).
- Blatt, M. R., Thiel, G. & Trentham, D. R. *Nature* **346**, 766-769 (1990).
- Kaplan, J. A. *Rev. Physiol.* **52**, 897-914 (1990).
- Mission, L., Taylor, C. W. & Berridge, M. J. *Nature* **352**, 241-244 (1991).
- Meyer, T., Holowka, D. & Stryer, L. *Science* **240**, 653-656 (1988).
- Minta, A., Kao, J. P. Y. & Tsien, R. Y. *J. Biol. Chem.* **264**, 8171-8178 (1989).
- Iino, M. *J. gen. Physiol.* **98**, 681-698 (1991).
- Iino, M. *J. gen. Physiol.* **95**, 1103-1122 (1990).
- Bezprozvanny, I., Watras, J. & Ehrlich, B. E. *Nature* **351**, 751-754 (1991).

- Fleischer, S. & Inui, M. A. *Rev. Biophys. biophys. Chem.* **18**, 333-364 (1989).
- Irvine, R. F. *FEBS Lett.* **263**, 5-9 (1990).
- Finch, E. A., Turner, T. & Goldin, S. M. *Science* **252**, 443-446 (1991).
- Watras, J., Bezprozvanny, I. & Ehrlich, B. E. *J. Neurosci.* **11**, 3239-3245 (1991).
- Berridge, M. J. *J. Biol. Chem.* **265**, 9583-9586 (1990).
- Meyer, T. *Cell* **64**, 675-678 (1991).
- Jaffe, L. F. *Proc. natn. Acad. Sci. U.S.A.* **88**, 9883-9887 (1991).
- Kasai, H. & Augustine, G. J. *Nature* **348**, 735-738 (1990).
- Lechleiter, J., Girard, S., Peralta, E. & Clapham, D. *Science* **252**, 123-126 (1991).
- Iino, M. *J. gen. Physiol.* **94**, 363-383 (1989).
- Downes, C. P., Mussat, M. C. & Michell, R. H. *Biochem. J.* **203**, 169-177 (1982).

ACKNOWLEDGEMENTS. We thank T. Yamazawa for help with the InsP_3 -binding protein assay. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture, Japan.

Control of CFTR chloride conductance by ATP levels through non-hydrolytic binding

P. M. Quinton & M. M. Reddy

Division of Biomedical Sciences, University of California, Riverside, California 92521, USA

SITE-SPECIFIC mutation¹ and membrane reconstitution² experiments provide compelling evidence that the product of the gene which is at fault in the disease cystic fibrosis, termed the cystic fibrosis transmembrane conductance regulator (CFTR)³, is a small-conductance chloride channel activated by phosphorylation⁴. As transport of chloride ions is passive, the predicted presence of two nucleotide-binding domains in CFTR seems as puzzling as a report⁵ that ATP hydrolysis is essential to activate the channel. We now find that in the sweat duct, which expresses high levels of CFTR⁶ and has a very high Cl⁻ conductance⁷, intracellular concentrations of ATP must be about normal (5 mM)⁸ for activation of this conductance, apparently by a non-hydrolytic, perhaps allosteric, mechanism. This passive dependence on ATP should mean that even a modest depletion of cell energy levels will significantly lower the energy demands of electrolyte transport by decreasing chloride conductance. We believe this direct coupling between cellular ATP levels and chloride channel activity is an adaptive mechanism to protect the tissue from damage resulting from excessive energy depletion⁹.

We initially suspected that chloride conductance (G_{Cl}) might be directly dependent on cell energy levels because metabolic poisoning of the isolated microperfused sweat duct with inhibitors of oxidative phosphorylation and/or glycolysis (2-deoxy-D-glucose (Fig. 1), iodoacetate, cyanide, and carbonyl cyanide *m*-chlorophenylhydrazone (data not shown) initially led to a surprising fall in transepithelial conductance (G_t). Significant reductions in cellular ATP are inevitably associated with tissue deterioration, which in epithelia is ultimately associated with a loss of barrier properties and an accompanying increase in G_t . We can be sure that the decrease in G_t is due to a specific decrease in G_{Cl} because (1) more than 80% of G_t is due to G_{Cl} in the sweat duct; (2) only minimal decreases in G_t occurred in the absence of Cl⁻ in the medium; and (3) ducts in cystic fibrosis patients are unresponsive to poisoning. To show that G_{Cl} is directly dependent on ATP, we manipulated cytoplasmic ATP concentrations after selectively permeabilizing the perfused duct. We found that the pore-forming agent α -toxin permeabilized only the basolateral membrane of the duct epithelial cells. The α -toxin pores reportedly pass molecules of $M_r \geq 3,500$ (ref. 10). When the basolateral surface was exposed to α -toxin, G_t characteristically decreased significantly during the next 10–20 min (Fig. 2a). After treatment with α -toxin, duct cells stained positively with trypan blue, which demonstrates permeation, and although they are normally impermeable, cyclic AMP together with 5 mM ATP caused a reversible increase in G_t and transepithelial voltage (V_t). Perfusion of the lumen with 0.1 mM ATP had no detectable effect on G_t or V_t . Two cystic fibrosis ducts treated similarly showed no detectable increase in G_t (data not shown).

Plots of G_t and V_t response against ATP concentration in the range 0–10 mM show that the response to 1 mM ATP was only ~20% of that to 5 mM ATP (Fig. 2). Assuming G_{Cl} to be maximal at the highest ATP concentration (10 mM) leads to an estimated EC_{50} (50% effective concentration) of more than 3 mM, which is about a 1,000 times higher than the K_m of ATP with the cAMP-dependent protein kinase PKA (3.1 μ M; ref. 11). The much higher concentration of ATP required to activate G_{Cl} here is not consistent with phosphorylation as the sole mechanism for regulating G_{Cl} . Recently, peptides made by recombinant

synthesis (corresponding to amino-acid residues 426–588; ref. 12) and by peptide synthesis (amino acids 450–516)¹³ to represent the region of the first nucleotide-binding fold of CFTR were found to bind ATP with relatively low affinity ($K_m = 5$ mM (ref. 12) and $K_m = 300$ μ M (ref. 13), respectively). ATP was not hydrolysed by either peptide. But as phosphorylation is a necessary but insufficient criterion to activate G_{Cl} and as stoichiometric hydrolysis of ATP is not required for passive Cl⁻ transport through the CFTR Cl⁻ channel, these results indicate that ATP could be acting allosterically in regulating G_{Cl} . We investigated whether non-hydrolytic binding of ATP is a conditional requirement for channel activation by testing the ability of non-hydrolysable and poorly hydrolysable ATP analogues to activate G_{Cl} . We found that, if phosphorylation by PKA is supported by maintaining cAMP and low concentrations of ATP, then both non-hydrolysable and hydrolysable analogues (Figs 3 and 4) can activate G_{Cl} , but under the same conditions AMP was clearly inhibitory.

A recent investigation⁵ of G_{Cl} in CFTR-transfected cell lines concluded that in addition to PKA-dependent phosphorylation, activation of G_{Cl} is also directly dependent upon ATP hydrolysis, with an EC_{50} of about 270 μ M. Our results in an organ tissue differ substantially in two ways. First, our apparent EC_{50} is at least 10-fold higher (>3 mM), and second, as non- and poorly hydrolysable ATP analogues also effectively increased G_{Cl} , ATP hydrolysis does not seem to be required beyond PKA phosphorylation. Other than the fact that the two systems are very different, we have no ready explanation as to why the G_{Cl} of a membrane patch from transfected cells should appear to be so much more sensitive to ATP than the intact apical membrane of microperfused sweat ducts. In contrast to the sweat duct, the G_{Cl} of transfected cells was insensitive to AMP, ADP, non-hydrolysable and poorly hydrolysable ATP analogues. This discrepancy might in part be due to the absence of the small amount of ATP needed to maintain phosphorylation at the time when the analogues were applied to the transfected cell mem-

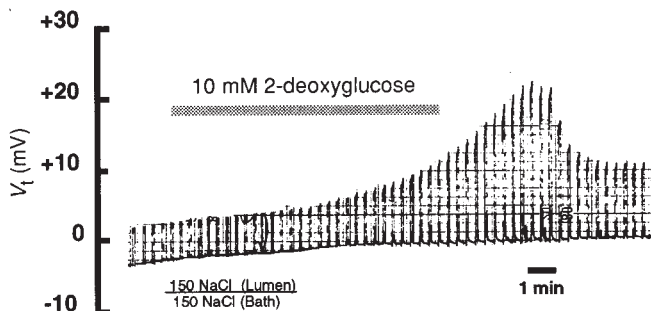


FIG. 1 Effect of metabolic energy depletion on the transepithelial potential (V_t) and conductance (G_t) of the microperfused sweat duct. Substitution of D-glucose with 2-deoxy-D-glucose, a non-metabolizable competitive inhibitor, removes the source of metabolic energy and depletes the energy supply of the tissue. As a result, V_t (mV), represented by the continuous trace, approaches 0 from normally negative values while the transepithelial resistance increases, as indicated by the increasing voltage deflections caused by constant current pulses of 50 nA for 1 s duration at 15-s intervals. As G_t is the reciprocal of resistance, conductance falls as a function of energy depletion. The effect is reversible if the inhibitor is removed before cell damage occurs.

METHODS. Sweat ducts were microdissected from sweat glands removed from human skin biopsies kept at 4 °C overnight in α -MEM medium (GIBCO) as described⁷. Duct segments were cannulated with a double-barrelled micropipette so that luminal perfusate could be changed and V_t monitored through one barrel while current pulses were applied through the second barrel. The duct was perfused with NaCl Ringer's solution, pH 7.4, containing (mM): NaCl (150), K⁺ (5), PO₄³⁻ (3.5), MgSO₄ (1.2), Ca²⁺ (1). The perfusion solution was identical, except that either 10 mM D-glucose or 2-deoxy-D-glucose was added as indicated. All experiments were at 35 ± 2 °C.

branes, although the different effects of poorly hydrolysable analogues like ATP- γ S remain unexplained. Still, non-hydrolytic regulation of ion channels is not without precedent. ATP-sensitive K^+ channels provide a good example of such control¹⁴.

Allosteric control of proteins by ATP is a well established phenomenon, especially in reactions that affect the energy level of the cell, such as those catalysed by phosphofructokinase and glucose-6-phosphate dehydrogenase¹⁵. The regulatory effect of

ATP on G_{Cl} implies that the heavy energy demands of active transport are coupled to the energy charge of the cell. Energy charge is defined as the ratio of the concentrations of $[ATP + \frac{1}{2}ADP]/[AMP + ADP + ATP]$. In pathways dedicated to energy production, a lower energy charge is stimulatory, whereas in pathways involved in energy consumption, a lower energy charge is inhibitory¹⁵. In this context, the activating effects of ATP and its analogues and the inhibitory effect of AMP on G_{Cl}

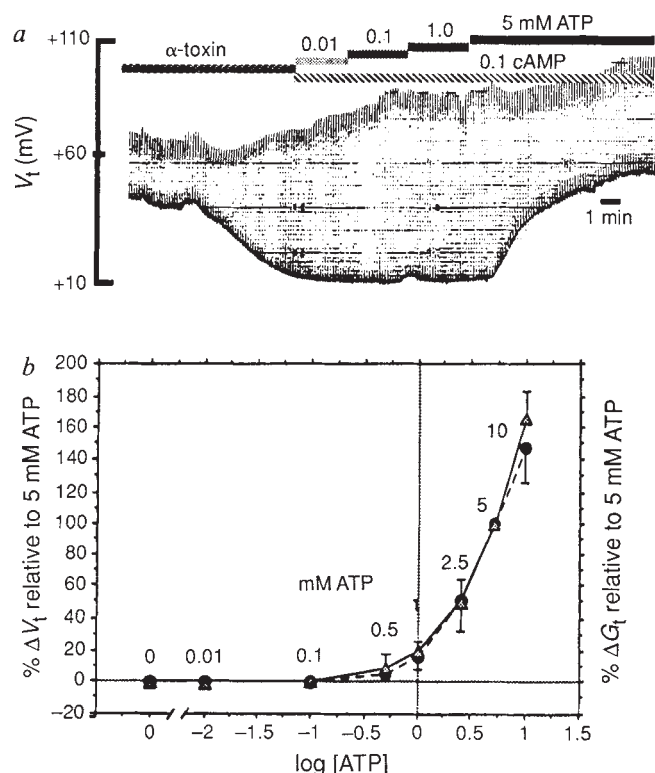


FIG. 2 Effect of α -toxin and increasing concentrations of ATP on transepithelial conductance (G_t) and Cl^- diffusion potentials (V_t) in the microperfused sweat duct. *a*, After exposing the basolateral membrane to 1,000 U ml⁻¹ of α -toxin from *Staphylococcus aureus*, V_t and G_t fall simultaneously. A single ion (Cl^-) gradient was established so that V_t predominantly reflects the Cl^- electromotive force and G_t predominantly reflects G_{Cl} across the apical membrane. In this example, the application of 0.1 mM cAMP had no effect until ATP concentration in the bath solution was increased to 5 mM, at which point V_t and G_t simultaneously increased, indicating activation of G_{Cl} . *b*, G_{Cl} does not show significant activation even in the presence of 0.1 mM cAMP until the concentration of ATP reaches approximately normal intracellular levels *in vivo* (4–7 mM¹¹). Except for 0.01 mM ATP ($n=3$), at least 6 measurements at each concentration were made in different sweat ducts ($n=6-16$).

METHODS. To establish the $[Cl^-]$ gradient, the lumen was perfused with the sodium conductance blocker amiloride (10 μ M) dissolved in NaCl Ringer's solution as defined in Fig. 1 while the duct was perfused with a Ringer's solution modified to reflect cytosolic composition containing (in mM): K^+ (145), gluconate (140), PO_4^{3-} (3.5), $MgSO_4$ (1.2), 260 μ M Ca^{2+} buffered with 2.0 mM EGTA (Sigma) to 80 nM free Ca^{2+} ; pH 6.8. When gluconate replaced Cl^- on both surfaces of the membrane, no response to ATP concentration plus ATP (5 mM) was observed (data not shown), demonstrating that no other ion contributed to the effect and that V_t and G_t are primarily a function of membrane G_{Cl} . The relative changes in transepithelial voltage (ΔV_t , filled circles) and conductance (ΔG_t , filled triangles) in response to ATP concentrations were calculated as percentages and normalized according to the general equation: $\Delta X = X_t/(X_5 - X_0)$ for each duct, where $X_t = V_t$ or G_t at a given ATP concentration, $X_5 = V_t$ or G_t at 5 mM ATP, and $X_0 = V_t$ or G_t at 0 mM ATP all in the same duct. The 100% response was taken at 5 mM ATP because that concentration most closely approximates the normal intracellular level, even though it is clear from the greater response to 10 mM ATP that the response at 5 mM is submaximal. Before normalization, the specific conductance of the tubule, G_t , was calculated from the cable equation⁷. All nucleotides were added as a free acid or monovalent salt.

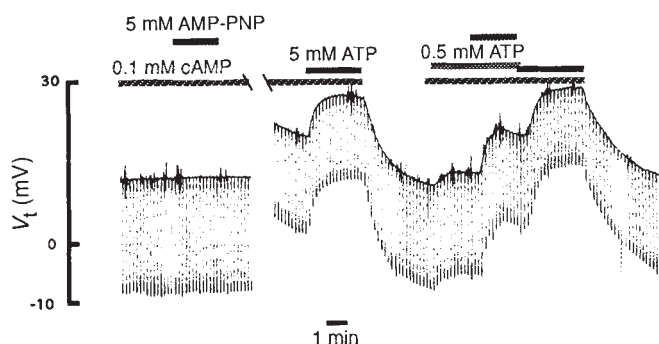


FIG. 3 ATP is required for effect of non-hydrolysable adenosine 5'-(β , γ -imino)triphosphate (AMP-PNP). Without ATP, 5 mM AMP-PNP had no detectable effect on V_t or G_t . But after adding a minimally activating concentration of ATP (0.5 mM), 5 mM AMP-PNP increased G_{Cl} as seen by the increase in V_t and G_t . cAMP (0.1 mM) was present as indicated and experimental conditions were as described in Fig. 2. The entire trace is from the same duct, but intervening manoeuvres were omitted as indicated by the break.

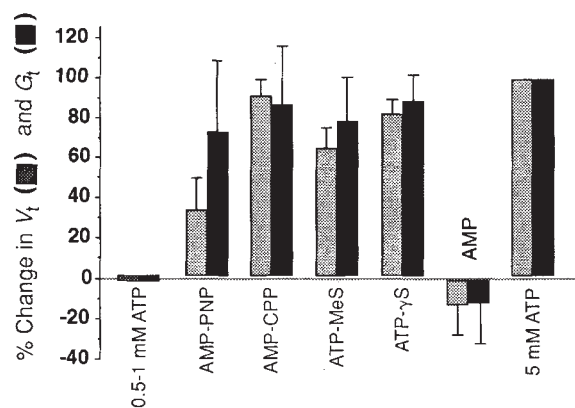


FIG. 4 Effect of nucleotides and ATP analogues on G_{Cl} . The ability of the different compounds to activate G_{Cl} relative to 5 mM ATP was determined by replacing ATP with a selected nucleotide at 5 mM concentration. cAMP (0.1 mM) was present in all cases. The relative effect on G_{Cl} was evaluated in terms of per cent changes in V_t (shaded bars) and G_t (filled bars) as described in Fig. 2, except that $X_t = V_t$ or G_t with 5 mM of each nucleotide plus 0.5 or 1.0 mM ATP and $X_0 = V_t$ or G_t at 0.5 or 1.0 mM ATP without any other nucleotide. Non-hydrolysable AMP-PNP had no effect in the absence of ATP, but in the presence of 1.0 mM ATP, the ATP analogue caused a significant increase in both V_t and G_t , showing that minimal ATP hydrolysis is required (probably to sustain phosphorylation), and also that non-hydrolytic binding is essential to activate the Cl^- conductance. The effect of poorly hydrolysable adenosine 5'-O-3-thiotriphosphate (ATP- γ S) was also enhanced by 1.0 mM ATP, even though significant activation of G_{Cl} by ATP- γ S did not require supplemental ATP. The other hydrolysable analogues, adenosine 5'-(α , β -methylene) triphosphate (AMP-CPP) and 2-methylthioadenosine triphosphate (ATP-MeS), were also capable of activating G_{Cl} in the absence of ATP (data not shown). AMP inhibited G_{Cl} activation, but although ADP was expected to be inhibitory it seemed to activate G_{Cl} (data not shown). But this effect was probably due to ATP as the response was completely blocked by diadenosine tetraphosphate, a specific inhibitor of adenylate kinase, which converts 2 ADP $\rightarrow$ ATP + AMP¹⁷. Measurements for each nucleotide were made in at least 3 different sweat ducts ($n=3-8$).

(Fig. 4) both provide strong support for such a regulatory function.

We interpret our findings to indicate that G_{Cl} , possibly via the first nucleotide-binding fold in CFTR, cannot be activated unless sufficient ATP is available to support active transport without compromising other ATP-dependent functions essential to cell survival. Teleologically, the mechanism would protect the cell from excessive energy depletion which would compromise the maintenance of cell volume, pH, and intracellular Ca^{2+} concentration, for example. Implicit in this conclusion is that a defective nucleotide-binding fold in CFTR may render this channel protein incapable of 'sensing' when intracellular ATP levels are high enough and so misdirect it to remain closed⁹. Recombinant and synthetic peptides corresponding to the first

nucleotide-binding fold with and without the mutant phenylalanine deletion $\Delta F508$ bind equally well with ATP^{12,13,16}, but this may not detract from our argument because defective binding might only be expressed in the environment of the whole protein *in situ*: this supported by the fact that the $\Delta F508$ peptide loses its ATP-binding capacity in 4 M urea denaturant but the wild-type peptide does not¹⁶.

To our knowledge, this report presents the first direct evidence in an epithelium that ATP modulates its own consumption by controlling the rate of active transport through feedback regulation of a passive component, Cl^- conductance. These results raise the question of whether defects in the nucleotide-binding folds in CFTR affect the cell's ability to sense its own energy levels for regulating ion transport in cystic fibrosis. □

Received 13 July; accepted 22 September 1992.

1. Anderson, M. P. *et al.* *Science* **253**, 202–205 (1991).
2. Bear, C. E. *et al.* *Cell* **68**, 809–818 (1992).
3. Riordan, J. R. *et al.* *Science* **245**, 1066–1072 (1989).
4. Tabcharani, J. A., Chang, X. B., Riordan, J. R. & Hanrahan, J. W. *Nature* **352**, 628–631 (1991).
5. Anderson, M. P. *et al.* *Cell* **67**, 775–784 (1991).
6. Kartner, N. *et al.* *Nature Genet.* (in the press).
7. Quinton, P. M. *Am. J. Physiol.* **251**, C649–C652 (1986).
8. Williamson, J. R. & Corkey, B. E. in *Meth. Enzym.* **13**, 434–513 (1969).
9. Quinton, P. M. & Reddy, M. M. *Adv. exp. Med.* **290**, 159–172 (1991).
10. Fussie, R. *et al.* *J. Cell Biol.* **91**, 83–94 (1991).
11. Flockhart, D. A., Friest, W., Hoppe, J., Lincoln, T. M. & Corbin, J. D. *Eur. J. Biochem.* **140**, 289–295 (1984).

12. Hartman, J. *et al.* *J. Biol. Chem.* **267**, 6455–6458 (1992).
13. Thomas, P. J., Shenbagamurthi, P., Ysern, X. & Pedersen, P. L. *Science* **251**, 555–557 (1991).
14. Ashcroft, S. J. H. & Ashcroft, F. M. *Cell Signal* **2**, 197–214 (1990).
15. Atkinson, D. E. *Metabolism and its Regulation* (Academic, San Francisco, 1977).
16. Thomas, P. J., Shenbagamurthi, P., Sondek, J., Hulihan, J. M. & Pedersen, P. L. *J. Biol. Chem.* **267**, 5727–5730 (1992).
17. Reinstein, J. *et al.* *Biochemistry* **29**, 7440–7450 (1990).

ACKNOWLEDGEMENTS. We thank K. Taylor for technical assistance. This work was supported by grants from the National Cystic Fibrosis Foundation (USA), the USPHS-NIH and by a gift from Texaco Corporation.

The protein kinase A-regulated cardiac Cl^- channel resembles the cystic fibrosis transmembrane conductance regulator

Georg Nagel, Tzyh-Chang Hwang, Kent L. Nastiuk*, Angus C. Nairn* & David C. Gadsby†

Laboratory of Cardiac/Membrane Physiology, and *Laboratory of Molecular and Cellular Neuroscience, The Rockefeller University, 1230 York Avenue, New York, New York 10021-6399, USA

STIMULATION of β -adrenoceptors in cardiac ventricular myocytes activates a strong chloride ion conductance^{1–5} as a result of phosphorylation by cyclic AMP-dependent protein kinase (PKA)^{2,4}. This Cl^- conductance, which is time- and voltage-independent^{1–5}, counters^{2,5} the tendency of the simultaneously enhanced Ca^{2+} channel current to prolong the ventricular action potential. Using inside-out giant patches⁶ excised from guinea-pig myocytes, we show here that phosphorylation by the PKA catalytic subunit plus Mg-ATP elicits discrete Cl^- channel currents. In almost symmetrical Cl^- solutions (~150 mM), unitary current amplitude scales with membrane potential, and reverses sign near 0 mV, to yield a single channel conductance of ~12 pS. Opening of the phosphorylated channels requires hydrolysable nucleoside triphosphate, indicating that phosphorylation by PKA is necessary, but not sufficient, for channel activation. The properties of these PKA-regulated cardiac Cl^- channels are very similar, if not identical, to those of the cystic fibrosis transmembrane conductance regulator (CFTR)⁷, the epithelial cell Cl^- channel whose regulation is defective in patients with cystic fibrosis. The full cardiologic impact of these Cl^- channels and of their possible malfunction in patients with cystic fibrosis remains to be determined.

To characterize cardiac ventricular chloride channels regulated by PKA and ATP, membrane current was recorded from excised, inside-out giant patches (pipette tip diameters, 12–20 μ m) held at 0 mV in the presence of a large transmembrane

gradient of Cl^- ions; the pipette/extracellular chloride ion concentration ($[Cl^-]$) was 160 mM, and the bath/cytoplasmic $[Cl^-]$ concentration was 4 mM. Although an 8-min exposure of the cytoplasmic surface to 0.5 mM Mg-ATP alone did not alter patch current (Fig. 1a), the subsequent addition of 100 nM PKA catalytic subunit led within 70 s (mean latency, 27 ± 16 s, s.d., $n = 10$) to the stepwise appearance of an outward current, which in this instance reached ~5 pA. Small, discrete current jumps of ~0.4 pA, reflecting the opening and closing of individual ion channels, are clearly visible (see also Figs 1b, 2, 3). PKA plus Mg-ATP activated this current in 39 out of 50 patches tested. Typically, 2 to 5 (but occasionally up to 15) channels could be open simultaneously. For the same large 160/4 mM $[Cl^-]$ gradient, the single channel current amplitude at 0 mV averaged 0.36 ± 0.03 pA ($n = 32$), and was about the same whether the principal pipette/extracellular cation was Na^+ (0.37 ± 0.03 pA, $n = 25$) or *N*-methyl-D-glucamine (0.35 ± 0.01 pA, $n = 7$), or whether the principal bath/cytoplasmic cation was Cs^+ (0.35 ± 0.02 pA, $n = 14$) or *N*-methyl-D-glucamine (0.37 ± 0.02 pA, $n = 18$), which suggests that the Cl^- ions were carrying the current.

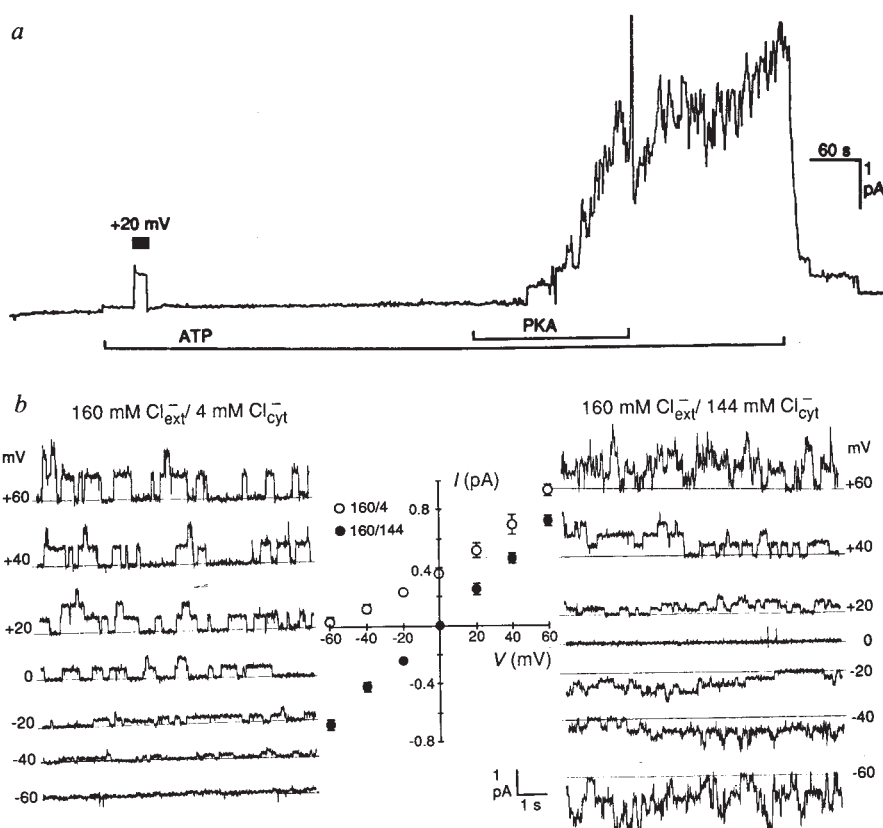
Further evidence of the chloride-selective nature of the PKA-activated channels was obtained by examining their current-voltage (*I*-*V*) relationship at different cytoplasmic Cl^- concentrations. Figure 1b shows (left) samples of current records at different holding potentials from a patch exposed to 160 mM Cl^- in the pipette and 4 mM Cl^- in the bath. With this steep $[Cl^-]$ gradient, the single-channel current remained outward, but its amplitude declined steadily along a shallow curve, on membrane hyperpolarization from +60 to -60 mV (Fig. 1b, open circles). In contrast, after raising the bath $[Cl^-]$ to 144 mM (Fig. 1b, right), the single-channel current amplitude from the same patch was more strongly voltage-dependent and reversed sign not far from -3 mV (Fig. 1b, filled circles), the estimated equilibrium potential for Cl^- ions (E_{Cl}). The single-channel *I*-*V* relationship was roughly linear under these approximately symmetrical $[Cl^-]$ conditions (Fig. 1b filled circles), yielding a single channel conductance of 12 ± 1 pS ($n = 3$), which is comparable to the 13 pS slope conductance obtained at large positive potentials for channels in cell-attached patches on intact myocytes stimulated by adrenaline (at ~35 °C; ref. 8). Also as in cell-attached patches⁸, the channel open probability is little affected by membrane potential over the 120 mV range tested (Fig. 1b).

† To whom correspondence should be addressed.

FIG. 1 Chloride channels in excised giant patches are activated by the catalytic subunit of PKA in the presence of Mg-ATP. *a*, Record of patch current at 0 mV showing responses to a voltage step (solid black bar) to +20 mV (indicating $\sim 25 \text{ G}\Omega$ seal resistance) and to exposures to Mg-ATP, with and without 100 nM PKA ($3.5 \mu\text{g ml}^{-1}$) as indicated, in the complete absence of Na^+ , K^+ or Ca^{2+} ions. Pipette $[\text{Cl}^-]$ was 160 mM, bath $[\text{Cl}^-]$ was 4 mM, and *N*-methyl-D-glucamine was the principal cation in both solutions, at 150 and 140 mM, respectively. The pipette solution also contained 5 mM Cs^+ . The Mg-ATP concentration was 500 μM in the absence of PKA, but 50 μM in its presence. Withdrawal of ATP caused rapid closure of the channels. Pipette tip diameter, 17 μm . *b*, Dependence of unitary current amplitude on voltage and on $[\text{Cl}^-]$ gradient. Channels were activated by PKA and 200 μM ATP, with 160 mM pipette/extracellular Cl^- (Cl_{ext}^-) and 4 mM bath/cytoplasmic Cl^- (Cl_{cyt}^-), and PKA was washed out with 500 μM ATP solution. The patch voltage was then stepped for 20 s to the levels indicated beside the traces on the left. Cl_{cyt}^- was then raised to 144 mM and the voltage steps repeated to yield the traces on the right. Dashed lines indicate current levels with all channels closed. Mean unitary current amplitudes (I) measured from amplitude histograms of opening and closing events are shown plotted against patch membrane potential (V) in the middle panel. Pipette and bath solutions were as in panel *a*; pipette tip diameter, 18 μm .

METHODS. Guinea-pig ventricular myocytes were isolated with collagenase and stored in medium with high $[\text{K}^+]$, low $[\text{Ca}^{2+}]$ (as in ref. 4) for 8–24 h so that large sarcolemmal blebs would develop⁶.

(We presume that blebbing of the sarcolemma has no influence on the appearance of the PKA-activated Cl^- channel currents described here, because their characteristics correspond well with those of single channel currents in cell-attached patches⁸, and of whole-cell currents^{1–5} activated within seconds of exposing intact, acutely dissociated, myocytes to catecholamines.) Wide-tipped (12–20 μm diameter) borosilicate glass pipettes of $\sim 100 \text{ k}\Omega$ resistance were made hydrophobic at the tips by coating them with a mineral oil-parafilm mixture⁶, then sealed to the blebs with light suction. After a gigaohm seal was established, the membrane patch was excised and transferred to a continuously perfused, temperature-controlled (25 $^{\circ}\text{C}$) chamber, in which solutions could be exchanged in $\sim 1 \text{ s}$ (ref. 6) by switching manual (Hamilton, Reno) or computer-controlled electric (General Valve, NJ) valves. Pipette and bath solutions, and the 0 mV holding potential, were designed to inhibit currents through Ca^{2+} , Na^+ , and K^+ channels, and the $\text{Na}^+/\text{Ca}^{2+}$ exchanger. The pipette solution usually con-



tained (in mM): 145 *N*-methyl-D-glucamine (NMG)-Cl, 5 CsCl (or, instead, 145 NaCl, 5.4 KCl), 2 BaCl_2 , 2.3 MgCl_2 , 0.5 CdCl_2 , 10 HEPES, pH 7.4 with NMG (or NaOH), and the bath solution usually contained: 140 NMG-aspartate, 10 EGTA, 10 HEPES, 2 MgCl_2 , 20 TEA-OH (to pH 7.4 with Tris). Na^+/K^+ pump current, activated by a brief exposure to bath Na^+ plus Mg-ATP²⁵, provided a convenient indicator of inside-out patch (rather than vesicle) formation. GTP, ADP and AMP-PNP were added as Tris, Na^+ , and Li^+ salts, respectively. The PKA catalytic subunit was prepared as described²⁶ and dialysed into (in mM) 10 EGTA, 10 HEPES, 20 TEA-Cl, 2 MgCl_2 , 85 aspartic acid (to pH 7.4 with $\sim 120 \text{ mM}$ CsOH), to a protein concentration of 0.7 mg ml^{-1} . Patch current was recorded with a LIST EPC-7 amplifier, stored on video tape, filtered at 100 Hz and digitized at 200 Hz off-line (in *b*; 7 Hz and 20 Hz, respectively, for other figures), and then analysed with programs written in ASYST.

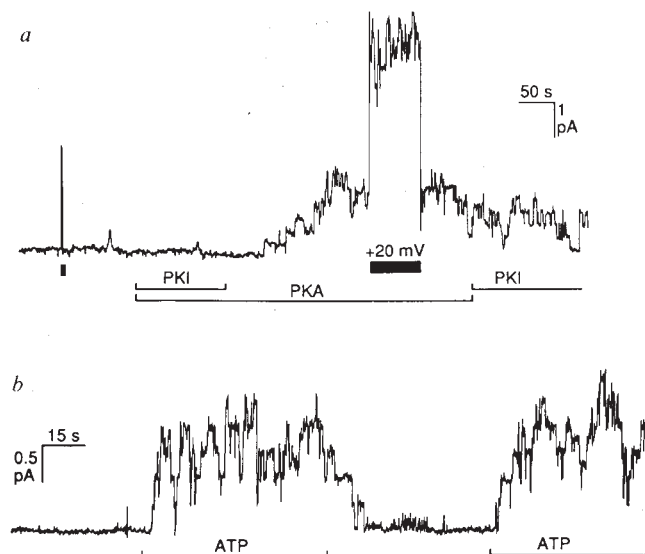
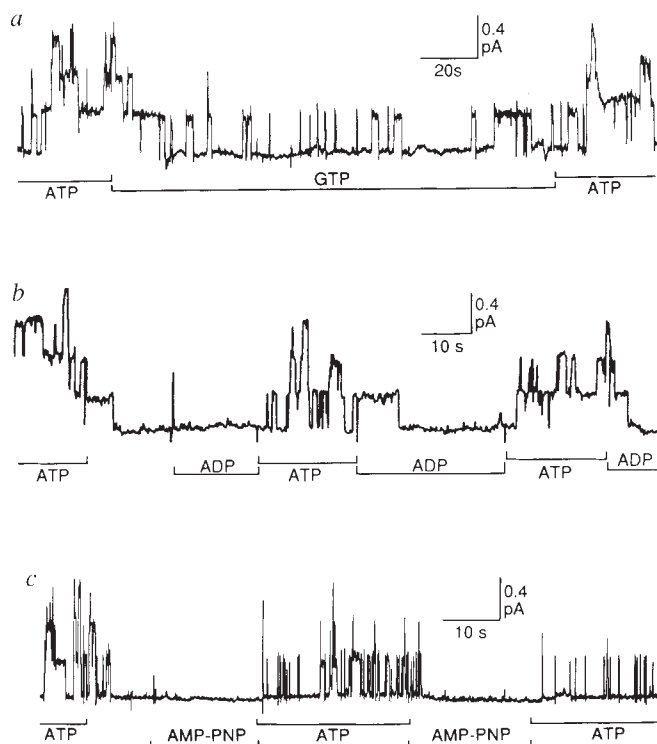


FIG. 2 Phosphorylated chloride channels need ATP before they can open. *a*, 100 μM PKI peptide (5–24-amide⁹) prevented channel activation by 100 nM PKA but failed to close the phosphorylated channels after withdrawal of PKA; 500 μM Mg-ATP was present throughout. Black bars beneath the current trace mark voltage steps to +20 mV, indicating a seal resistance of $\sim 7 \text{ G}\Omega$. The holding potential was 0 mV; the pipette solution was the NaCl plus KCl version; the bath solution was as in Fig. 1*a*, except the principal cation was Cs; pipette tip diameter was 12 μm . *b*, Reversible activation of phosphorylated Cl^- channels by ATP. The current record begins $\sim 30 \text{ s}$ after withdrawal of PKA and ATP had closed all channels. Readdition of 500 μM Mg-ATP (as indicated by the lower lines) re-opened the channels. This effect of ATP was still observed in this patch even 15 min after the removal of PKA. Holding potential was 0 mV; pipette and bath solutions were as for Fig. 1*a*; pipette tip diameter, 18 μm .

FIG. 3 Hydrolysable nucleoside triphosphates are required to open PKA-phosphorylated Cl⁻ channels. *a*, Partial restoration of Cl⁻ channel activity by GTP. The lower lines indicate when 500 μ M ATP was exchanged for 500 μ M GTP, and vice versa. Pipette tip diameter, 17 μ m. *b*, Replacing ATP with ADP fails to sustain Cl⁻ channel activity. The lower lines indicate the withdrawal and/or addition of 500 μ M ATP and/or 500 μ M ADP and vice versa (switching artefacts from electric valves are visible). Pipette tip diameter, 18 μ m. *c*, AMP-PNP, a non-hydrolysable ATP analogue, failed to re-open the Cl⁻ channels closed by removal of ATP. Lower lines indicate the withdrawal and/or addition of 500 μ M ATP and/or 500 μ M AMP-PNP. Pipette tip diameter, 18 μ m. For all three panels, the holding potential was 0 mV; pipette and bath solutions were as for Fig. 1*a*, except as noted; time calibration is 20 s for *a*, but 10 s for *b* and *c*. The progressive shortening of channel open times (Fig. 3*c*) was occasionally observed in other patches, and possibly underlies the previously reported decay or 'rundown' of macroscopic Cl⁻ current flowing through overexpressed CFTR channels¹³.



In dialysed myocytes with 20 mM internal Cl⁻ and 150 mM external Cl⁻ at 36 °C (E_{Cl} near -50 mV), isoprenaline or forskolin typically elicit whole-cell Cl⁻ currents of ~200 pA at 0 mV (refs 1-5). Figure 1*b* indicates that the single channel current amplitude would be ~0.4 pA under those conditions, requiring ~500 open channels per cell, which, taking an open probability of ~0.5 (as in Fig. 1*b*), gives a cell population of ~1,000 channels. For a cell-surface area of 18,000 μ m² (assuming a specific membrane capacitance of 1 μ F cm⁻² and a cell capacitance of 180 pF; ref. 4), average channel density would be ~0.06 μ m⁻² (compare with ref. 8), predicting ~11 channels in a 16- μ m diameter circular patch of membrane, which is in the range we observe.

Activation of these cardiac Cl⁻ channels requires PKA-mediated phosphorylation because the channels are not activated by Mg-ATP alone (Fig. 1*a*), by PKA catalytic subunit in the absence of Mg-ATP (not shown), or by PKA plus Mg-ATP in the presence of the specific PKA inhibitor peptide, PKI (5-24-amide⁹; Fig. 2*a*). In the latter case, withdrawal of PKI from the solution containing PKA catalytic subunit leads to the appearance of single chloride-channel currents after the previously observed latency (compare Figs 1*a* and 2*a*). Moreover, once phosphorylated, the channels remain active and can survive many minutes (≥ 15 min) of continuous washing to ensure complete removal of the PKA catalytic subunit, either without (see Fig. 1*a*, *b*, for example), or with (Fig. 2*a*) PKI present to prevent further kinase activity.

The persistent nature of the Cl⁻-channel currents in excised giant patches contrasts with the rapid decay of the corresponding whole-cell current on withdrawal of agonists or introduction of PKI to the cell interior²⁻⁵, and implies that dephosphorylation proceeds slowly in the excised patches. Surprisingly, however, the phosphorylated channels close promptly when ATP is withdrawn (Fig. 1*a*). This channel closure cannot be attributed to dephosphorylation, because the effect of ATP is readily and rapidly reversible in the complete absence of PKA; reapplication of 500 μ M ATP promptly reopens the channels (Fig. 2*b*). Indeed, the channels must have remained phosphorylated while deprived of ATP, because ATP alone fails to open non-phos-

phorylated channels (Fig. 1*a*). Analysis of channel open probability, at ATP concentrations from 10 μ M to 2 mM, suggests that ATP acts with relatively low affinity, with half-maximal effects requiring > 50 μ M ATP (data not shown).

To investigate this role of ATP further, we examined the effects of hydrolysable and non-hydrolysable analogues (Fig. 3). Only hydrolysable nucleoside triphosphates opened phosphorylated cardiac Cl⁻ channels. Thus, 500 μ M GTP was able to sustain channel activity (Fig. 3*a*), although at a lower open probability than seen with 500 μ M ATP in the same patch. Neither 500 μ M ADP (Fig. 3*b*), which has no γ -phosphate, nor 500 μ M AMP-PNP (Fig. 3*c*), which cannot donate its γ -phosphate, were able to open phosphorylated Cl⁻ channels that had been closed by withdrawal of ATP.

Our results show that there is a close biophysical and biochemical resemblance between the native PKA-regulated cardiac Cl⁻ channel and CFTR channels overexpressed in a variety of cell types: both show approximately voltage-independent gating¹⁰⁻¹², low single-channel conductance¹⁰⁻¹⁴, and linear unitary *I*-*V* relationships in symmetrical [Cl⁻]^{11,13}; both require phosphorylation by PKA^{10,11,13,15}, and both need, in addition, high micromolar levels of a hydrolysable nucleoside triphosphate to allow normal gating¹³. To support their identity further, our preliminary measurements of shifts of reversal potentials of whole-cell isoprenaline-induced currents upon replacing external Cl⁻ with test anion (with constant internal [Cl⁻] (125 mM)) suggest a permeability sequence Br⁻ > Cl⁻ $\geq$ I⁻ > F⁻ (T.-C.H., A. Dousmanis and D.C.G., unpublished observations; compare with ref. 16), which is like that of overexpressed CFTR¹⁷. Although the PKA-regulated Cl⁻ conductance in myocytes dialysed (access resistance 3-12 M Ω) with pipette solution containing 10 mM EGTA and 10 mM HEPES was diminished^{2,3} by high concentrations of DNDS and SITS (4,4'-dinitrostilbene-2,2'-disulphonic acid and 4-acetamido-4'-isothiocyanatostilbene-2,2'-disulphonic acid), these agents had no effect in myocytes dialysed with pipettes containing 50 mM EGTA and 40 mM HEPES², or in vigorously dialysed myocytes (access resistance 1-4 M Ω)¹⁸, indicating that the cardiac Cl⁻ channel, like CFTR¹⁹, is not directly blocked by stilbenes.

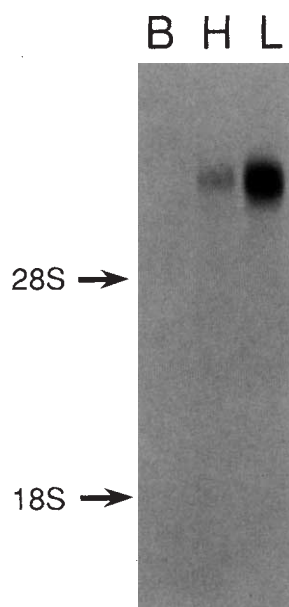


FIG. 4 Northern blot analysis of RNA from guinea-pig brain (B), heart (H) and lung (L). Total RNA was isolated by guanidinium thiocyanate extraction and sedimentation in CsCl²⁷. Total RNA (40 µg) from each tissue was denatured and separated on a formaldehyde-agarose gel²⁸. The gel was stained with ethidium bromide, photographed to confirm that 28S and 18S ribosomal RNAs were intact (arrowed), and the RNA was transferred to a nylon membrane (Zeta-probe, Bio-Rad). The filter was hybridized according to a modification²⁹ of the method described in ref. 30. Briefly, the filter was hybridized in 1 M Na⁺, 7% SDS, 10% polyethylene glycol (*M*_w 8,000) and 50% formamide at 65 °C with the ³²P-labelled antisense riboprobe, made by *in vitro* transcription of a 4.7-kb human cDNA⁷ clone containing the entire CFTR coding sequence, washed with 0.1 × SSC, 0.1% SDS at 65 °C and exposed for 3 days at -70 °C with intensifying screens.

Northern analysis indicates that CFTR messenger RNA is expressed in the guinea-pig heart. Figure 4 shows bands of ~6.9 kb, which are similar in size to CFTR mRNA⁷, in samples of total RNA from heart and lung, but not brain, when hybridized with a full-length antisense riboprobe derived from human CFTR complementary DNA. In contrast, the complementary sense-strand probe did not detectably hybridize to any guinea-pig RNAs tested (not shown). The strong implication that PKA-regulated cardiac Cl⁻ channels are, in fact, CFTR is in keeping with a recent demonstration²⁰ that expression of CFTR in non-epithelial cells is more widespread than expected.

Upon β-adrenoceptor stimulation, the normal steep chloride gradient drives a large outward, repolarizing Cl⁻ current through these channels at positive potentials¹⁻⁵, reducing the duration of the guinea-pig cardiac action potential⁵ by overcoming the tendency of the simultaneously enhanced Ca²⁺ current to prolong it^{2,5}. Prolongation of the cardiac action potential lengthens the refractory period and, as in the case of the long QT syndrome²¹, could predispose the heart to arrhythmias. The electrocardiological literature on cystic fibrosis patients is sparse, but ventricular dysfunction during stress²² and arrhythmias^{23,24} have been reported. If confirmed, one explanation is that PKA-mediated regulation of Cl⁻ channels, but not of Ca²⁺ channels, is defective in the hearts of cystic fibrosis patients, so that sympathetic stress might tend to prolong, rather than shorten, their cardiac action potentials. Alternatively, these Cl⁻ channels might contribute little to the human cardiac action potential, with the reported dysfunction being attributed to *cor pulmonale* related to cystic fibrosis. A clear answer will require an increased understanding of the expression and mechanisms of regulation of cardiac Cl⁻ channels in cystic fibrosis patients. □

Received 27 July; accepted 21 September 1992.

- Harvey, R. D. & Hume, J. R. *Science* **244**, 983-985 (1989).
- Bahinski, A., Nairn, A. C., Greengard, P. & Gadsby, D. C. *Nature* **340**, 718-721 (1989).
- Matsuoka, S., Ehara, T. & Noma, A. *J. Physiol., Lond.* **425**, 579-598 (1990).
- Hwang, T.-C., Horie, M., Nairn, A. C. & Gadsby, D. C. *J. gen. Physiol.* **99**, 465-489 (1992).
- Harvey, R. D., Clark, C. D. & Hume, J. R. *J. gen. Physiol.* **95**, 1077-1102 (1990).
- Hilgemann, D. W. *Nature* **344**, 242-245 (1990).
- Riordan, J. R. *et al. Science* **245**, 1066-1073 (1989).
- Ehara, T. & Ishihara, K. *Nature* **347**, 284-286 (1990).
- Cheng, H.-C. *et al. J. biol. Chem.* **261**, 989-992 (1986).
- Tabcharani, J. A., Chang, X.-B., Riordan, J. R. & Hanrahan, J. W. *Nature* **352**, 628-631 (1991).
- Berger, H. B. *et al. J. clin. Invest.* **88**, 1422-1431 (1991).
- Kartner, N. *et al. Cell* **64**, 681-691 (1991).
- Anderson, M. P. *et al. Cell* **67**, 775-784 (1991).
- Anderson, M. P., Rich, D. R., Gregory, R. J., Smith, A. E. & Welsh, M. J. *Science* **251**, 679-682 (1991).
- Cheng, S. H. *et al. Cell* **66**, 1027-1036 (1991).
- Overholt, J. L. & Harvey, R. D. *Biophys. J.* **61**, A442 (1992).
- Anderson, M. P. *et al. Science* **253**, 202-205 (1991).
- Hwang, T.-C., Horie, M., Dousmanis, A. G. & Gadsby, D. C. *J. gen. Physiol.* **100**, 69a (1992).
- Cliff, W. H., Schoumacher, R. A. & Frizzell, R. A. *Am. J. Physiol.* **262**, C1154-C1160 (1992).
- Yoshimura, K. *et al. Nucleic Acids Res.* **19**, 5417-5423 (1991).
- Vincent, G. M., Abildskov, J. Z. & Burgess, M. J. *Prog. Cardiovasc. Dis.* **16**, 523-530 (1974).
- Benson, L. N. *et al. Am. Rev. Respir. Dis.* **130**, 987-992 (1984).
- Chéron, G., Paradis, K., Stéru, D., Demay, G. & Lenoir, G. *Acta paediatr. scand.* **73**, 697-700 (1984).
- Sullivan, M. M., Moss, R. B., Hindi, R. D. & Lewiston, N. J. *Chest* **90**, 239-242 (1986).
- Hilgemann, D. W., Nagel, G. A. & Gadsby, D. C. In *The Sodium Pump: Recent Developments* (eds Kaplan, J. H. & De Weer, P.) 543-547 (Rockefeller Univ. Press, New York, 1991).
- Kaczmarek, L. Y. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 7487-7491 (1980).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1989).
- Amasino, R. M. *Analyt. Biochem.* **152**, 304-307 (1986).
- Church, G. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1991-1995 (1984).

ACKNOWLEDGEMENTS. We thank D. Hilgemann for teaching us his giant patch technique, L.-C. Tsui and G. Cutting for CFTR cDNA, and P. Hoff for technical assistance. Supported by NIH, the New York Heart Association, and the Cystic Fibrosis Foundation.

A new tropomyosin essential for cytokinesis in the fission yeast *S. pombe*

Mohan K. Balasubramanian*, David M. Helfman† & Sean M. Hemmingsen‡§

* Department of Microbiology, University of Saskatchewan, Saskatoon, Saskatchewan S7N 0W0, Canada

† Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

‡ Plant Biotechnology Institute, National Research Council, Saskatoon, Saskatchewan, S7N 0W9, Canada

MUTATIONS in the *Schizosaccharomyces pombe cdc8* gene impair cytokinesis¹. Here we clone *cdc8*⁺ and find that it encodes a novel tropomyosin. Gene disruption results in lethal arrest of the cell cycle, but spore germination, cell growth, DNA replication and mitosis are all unaffected. Haploid *cdc8* gene disruptants are rescued by expression of a fibroblast tropomyosin complementary DNA. Immunofluorescence microscopy of wild type and *cdc8* gene disruptants indicates that *cdc8* tropomyosin is present in two distinct cellular distributions: in dispersed patches, and during cytokinesis as a transient medial band. Collectively these results indicate that *cdc8* tropomyosin has a specialized role which, we suggest, is to form part of the F-actin contractile ring at cytokinesis. These results establish the basis for further genetic studies of cytokinesis and of contractile protein function in *S. pombe*.

Models of cell-cycle regulation focus on the roles of proteins such as *cdc2* kinase and cyclins in controlling nuclear events of the cycle². The links between these nuclear events and the initiation and completion of cytokinesis are not known. In *S. pombe*, recessive, conditionally lethal mutations at the *cdc3*, *cdc4*, *cdc7*, *cdc8*, *cdc11*, *cdc12*, *cdc14*, *cdc15* and *cdc16* loci lead to defects in cytokinesis and cell separation^{1,3,4}.

Mutations in *cdc8* arrest the cycle at cytokinesis but do not impair DNA replication or mitosis¹. At 25 °C *cdc8*-110 cells

§ To whom correspondence should be addressed.



FIG. 1 Isolation and sequence of the *cdc8* gene. **a**, Rescue of *cdc8-110* mutation by subclones of the 9-kb clone RES3. Rescue +, cells grow at 37 °C with normal morphology; Rescue -, cells die at 37 °C with *cdc*⁻ phenotype, as elongated single cells. pMB806.1 was produced by inserting the *ura4*⁺ gene into the indicated *HindIII* site, thus disrupting the predicted coding region after the fourth codon. The extent of the two exons is indicated by a bold interrupted arrow. **b**, Complete nucleotide sequence of MB805.1 and predicted amino-acid sequence of the interrupted open reading frame. Intron consensus sequences have been underlined. Restriction enzyme digestion sites: H, *HindIII*; X, *XhoI*; K, *KpnI*; S, *SacI*.

METHODS. A *cdc8-110 leu1-32* strain was transformed¹⁹ with two independent *S. pombe* genomic libraries in pWH5 (ref. 20). Selection was applied first for leucine prototrophy and then for colony formation at 37 °C. Overlapping clones that rescued *cdc8-110* were isolated from both libraries. Subclones of pRES3 were tested for complementation. pMB805.1 was the smallest rescuing subclone. pBX, which does rescue *cdc8-110*, was confirmed by Southern analysis to carry an authentic continuous 2.6-kb genomic fragment that included the sequences present on pMB805.1. An integrated copy of *LEU2* was introduced in place of the H-K-H region of pBX and used to confirm linkage of the cloned gene to the *cdc8-110* mutation: no *cdc8-110* segregants were observed among 400 *Leu*⁺ colonies. A *S. pombe* cDNA library constructed in the plasmid vector pTZ19R was screened with radioactively labelled *cdc8* genomic sequences. Several cDNA clones were isolated, nucleotide sequence analysis of which confirmed the splice junction sequences predicted from the *cdc8* genomic sequence.

grow and divide, forming colonies. At 37 °C, the cycle is arrested and cells die with 2-4 nuclei and disorganized septum material without forming colonies¹.

S. pombe genomic clones permitting growth of *cdc8-110* at 37 °C were isolated (see Fig. 1 legend). Determination of the sequence of the smallest complementing subclone identified a potential coding region interrupted by an intron (Fig. 1). Analysis of cDNA clones confirmed the presence of the predicted intron (Fig. 1 legend). Genetic mapping of an integrated marker established that the cloned gene is *cdc8* or a locus very tightly linked to it. The predicted *cdc8* product is a 161-amino-acid protein, several features of which identify it as a tropomyosin:

(1) sequence similarity (Fig. 2); (2) high predicted⁵ α -helical content (98.5%)⁶; and (3) a repeating 7-residue pattern of non-polar and polar or charged residues, characteristic of coiled-coil α -helical proteins⁷.

Translation initiation at codon 384 (ATG), which is in the optimal context for initiation⁸, would result in a 161-amino-acid polypeptide. Sequences upstream of this do not resemble known tropomyosins. The electrophoretic migration of a fusion polypeptide and of authentic *cdc8* from *S. pombe* are also consistent with *cdc8* encoding a polypeptide of 161 amino acids (Fig. 3). *cdc8* encodes the smallest tropomyosin known and it would bind an integral number (4) of actins⁶. The next smallest, the

<i>cdc8</i>	1	MDKLREKINAAEAETDEAVARAEEAEAKLKEVELQLSLKEQEYESLSRKS	50
<i>TPM1</i>	1	MDKIREKLSNLKLEAESWQEKYEELKEKNKOLEQENVEKENQIKSLTVKN	50
<i>cdc8</i>	51	EAAESQLEEELEETKQLRKADNEDIQKTEAEQ.....	83
<i>TPM1</i>	51	QLEDEIEKLEAGLSDSKQTFQDNVEKENQIKSLTVKNHQLLEEIEKLEA	100
<i>cdc8</i>	84	LSRKVELLEELEETNDKLLRETTEKMRQT	112
<i>TPM1</i>	101	ELAESQKLSSEDSHLLQSNNDNFSSKNQQLLEEDLEESDTKLKETTEKLR	150
<i>cdc8</i>	113	DVKAHFERRVQSLERERDDMEQKLEEMTDKYTKVKAELDEVHQALEDL	161
<i>TPM1</i>	151	DLKADQLERRVAALAEQRFVWERKNEELTVKYEDAKKELDEIAASLENL	199

FIG. 2 Sequence of the predicted *S. pombe cdc8* gene product compared with *S. cerevisiae TPM1* tropomyosin⁹. A dot indicates an internal gap introduced in the *cdc8* sequence. With this alignment, 60 residues are identical in *cdc8* and *TPM1* products.

199-amino-acid *Saccharomyces cerevisiae* TPM1 tropomyosin⁹, shows the highest degree of similarity to *cdc8* (Fig. 2).

To determine the effect of *cdc8* gene disruption, a heterozygous diploid strain (*cdc8*⁺/*cdc8*::*ura4*⁺) was constructed in which the disrupted allele encodes only the four N-terminal amino acids (Fig. 1). Haploid spores carrying the disrupted allele germinate, but die as elongated single cells with a cell-

cycle arrest morphology indistinguishable from that of *cdc8*-110. Plasmid loss from rescued *cdc8* disruptants also produced this phenotype. Nuclear staining identified the presence of up to six nuclei in arrested cells. We conclude that *cdc8* gene function is essential for cytokinesis but is not required for spore germination, cell growth, DNA replication or mitosis.

Polyclonal antisera were raised against purified *cdc8* fusion protein synthesized in *Escherichia coli*. These sera recognize the 175-amino-acid fusion polypeptide antigen and *cdc8* tropomyosin, a polypeptide with a slightly lower apparent *M_r* and detectable in *S. pombe* total and heat-stable protein extracts (Fig. 3). Immunoreactive tropomyosin is present at similar levels in heat-stable protein extracts from *cdc8*⁺ or *cdc8*-110 cells, or *cdc8*::*ura4*⁺ cells carrying *cdc8*⁺ on a multicopy plasmid, pRES3. Immunoreactive material is not detectable in total or heat-stable extracts of *cdc8*::*ura4*⁺ cells which are rescued by expression of a rat fibroblast tropomyosin cDNA carried on pTM-2 (Fig. 3). The 284-amino-acid rat tropomyosin isoform encoded by this cDNA¹⁰ is not recognized by anti-*cdc8* serum.

Immunofluorescence microscopy of wild-type *S. pombe* cells reveals that anti-*cdc8* tropomyosin-reactive material is found in all cells as small distributed patches (Fig. 4). In many binucleate cells with well separated nuclei, immunoreactive material is also

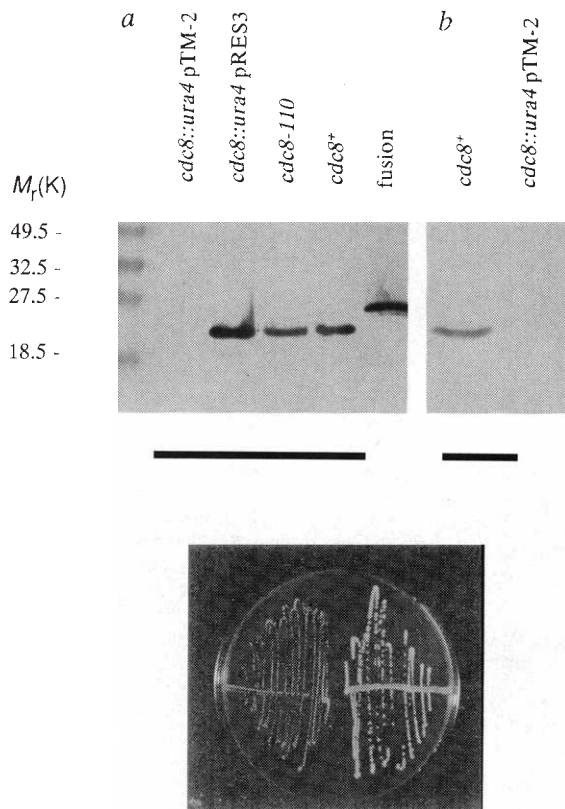


FIG. 3 Immunoblot analysis of *S. pombe* *cdc8* tropomyosin and colony formation by *cdc8*⁺ and *cdc8*::*ura4*⁺ pTM-2 cells. Bottom, photograph of the plate shows colony formation by *cdc8*::*ura4*⁺ cells rescued by pRES3 (right) or pTM-2 (left). Top, protein extracted from cells of the genotypes indicated was resolved by SDS-PAGE and immunoblotted with rabbit polyclonal anti-*cdc8* tropomyosin antiserum. *a*, Heat-stable protein probed with immune serum (1:1,500 dilution); *b*, total protein probed with immune serum (1:5,000 dilution) preadsorbed against total protein extracted from *cdc8*::*ura4*⁺ pTM-2 cells. Plasmid TM-2 carries an *adh* promoter-fibroblast tropomyosin isoform TM-2 cDNA¹⁰ construct (M.K.B., D.M.H. and S.M.H., manuscript in preparation). Tropomyosin migrates anomalously in SDS-PAGE⁹. In this gel system, *cdc8* tropomyosin, with a predicted *M_r* of 19,000, migrates as a 23–24K polypeptide.

METHODS. Antibody production: A gene encoding the first 16 amino acids of gene 10 of T7 bacteriophage²¹ fused to the C-terminal 159 amino acids of *cdc8* tropomyosin was expressed in *E. coli*. The gene product was purified almost to homogeneity¹¹ by fractionating a cleared homogenate by precipitation with ammonium sulphate, followed by heat treatment and ion-exchange chromatography. One antiserum was raised against this soluble material and a second after electrophoresis and recovery of the fusion polypeptide from a gel. A 1:1 mixture of these sera was used to probe the filter in *a*. The first serum (50 μ l) was preadsorbed against 60 mg of total protein extracted from *cdc8*::*ura4*⁺ pTM-2 cells. This serum was used to probe the filter in *b* and for indirect immunofluorescence microscopy. Total protein extracts: *S. pombe* cells were suspended in 50 mM Tris-HCl, 10 mM MgSO₄, 1 mM EDTA, 10% glycerol, 50 mM KCl, 1 mM dithiothreitol, pH7.5, containing pepstatin, leupeptin, aprotinin and chymostatin, each at 2 μ g ml⁻¹ and broken by passage through a French pressure cell. The homogenate was centrifuged at 20,000*g* for 20 min. For heat-stable protein extracts, total protein was made 1 M in NaCl and maintained at 95 °C for 10 min. Denatured protein was removed by centrifugation at 14,000*g* for 20 min. Immunoblots were developed with an alkaline phosphatase-conjugated secondary antibody and nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP).

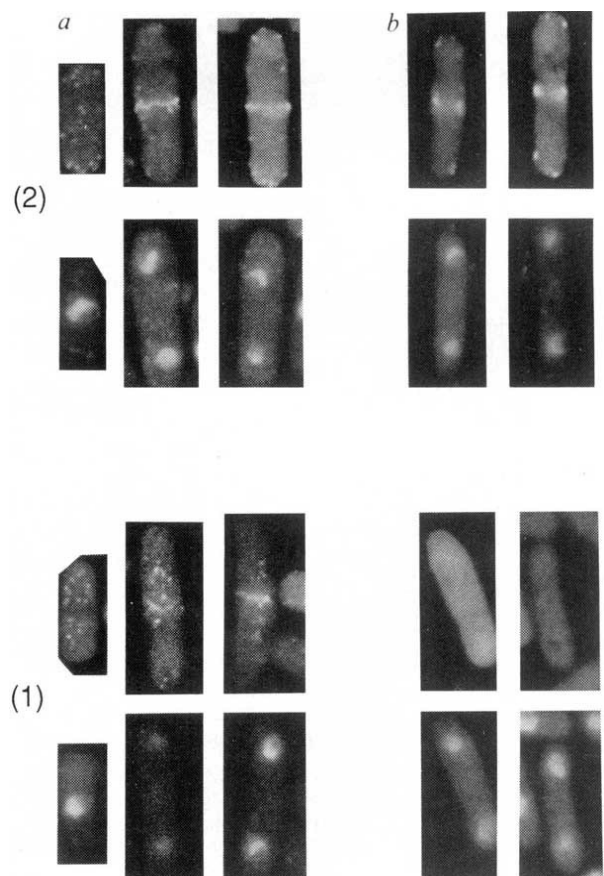


FIG. 4 Localization of *cdc8* tropomyosin and F-actin in *S. pombe* cells by indirect immunofluorescence microscopy: *a*, *cdc8*⁺ *leu1*-32 cells, or *b*, *cdc8*::*ura4*⁺ cells rescued by pTM-2, stained with (1) preadsorbed anti-*cdc8* tropomyosin (1:5,000 dilution) (upper images) and 4',6-diamidino-2-phenylindole (DAPI) (lower images), or (2) rhodamine-conjugated phalloidin (upper images) and DAPI (lower images).

METHODS. Cells were fixed with formaldehyde and stained with rhodamine-conjugated phalloidin¹² or fixed with formaldehyde and glutaraldehyde and processed for indirect immunofluorescence as described and mounted in medium containing DAPI²². Antibodies are as described in Fig. 3 legend.

apparent as a medial band. The antiserum used has been depleted against total protein extracted from *cdc8::ura4⁺* cells that carry pTM-2. This serum stains neither the patches nor the medial band in *cdc8::ura4⁺* pTM-2 cells and produces images that resemble those produced by preimmune serum (results not shown). It can be concluded that *cdc8* tropomyosin is present in wild-type cells both in the small distributed patches and in the transient medial band. Staining of *cdc8⁺* and *cdc8::ura4⁺* pTM-2 cells with serum that has not been depleted against *cdc8::ura4⁺* pTM-2 cell protein reveals that the medial band is absent from *cdc8::ura4⁺* pTM-2 cells, although the pattern of immunoreactive dots is still present in this strain (not shown). From this it can be concluded that if any other immunologically recognized tropomyosin isoforms are present in *S. pombe*, they do not contribute significantly to the medial band staining. A putative tropomyosin has been previously identified in *S. pombe*, although its relationship to those described here has not been determined¹¹.

Tropomyosin is an actin-binding protein⁷. We have compared the distribution of actin visualized by fluorescence microscopy to that of *cdc8* tropomyosin. The distribution that we observed for actin was similar to that for *cdc8* tropomyosin (Fig. 4). Actin stained as small distributed patches in all cells and as a medial band in cells with separated nuclei. Both patterns of staining persisted in *cdc8::ura4⁺* pTM-2 cells. The presence of actin in the medial plane may precede that of tropomyosin because actin staining was visible in cells in which daughter nuclei were relatively close together (not shown). Our results are similar to those reported previously for the distribution of actin in *S. pombe*¹²⁻¹⁴.

We have shown that tropomyosin has an essential role in *S. pombe*, but *S. cerevisiae* has a single tropomyosin gene which is non-essential for growth, cell division or sporulation⁹. The viability of *S. cerevisiae* in which this gene is disrupted may reflect a fundamental difference in the cell-cycle mechanisms of these two yeasts¹⁵. But other genes encoding contractile proteins, including an actin-like gene, *ACT2* (ref. 16), and a myosin gene, *MYO1* (ref. 17), have been implicated in cytokinesis. Thus, a second unidentified tropomyosin similar to the *S. pombe cdc8* tropomyosin may act in cytokinesis in *S. cerevisiae*. Cytokinesis in *S. pombe*, as in animal cells, is characterized by the presence of an F-actin contractile ring^{14,18}. We suggest that an essential role for *cdc8* tropomyosin is to form part of the F-actin contractile ring during cytokinesis. □

Zebrafish *pax[b]* is involved in the formation of the midbrain–hindbrain boundary

Stefan Krauss*, Malcolm Maden, Nigel Holder & Stephen W. Wilson

Developmental Biology Research Centre, Randall Institute, Division of Biomedical Sciences, King's College, 26–29 Drury Lane, London WC2B 5RL, UK

* Present address: ICRF, Developmental Biology Unit, Department of Zoology, South Parks Road, Oxford OX1 3PS, UK

AMONG the genes thought to be involved in patterning the nervous system are a family of developmentally regulated paired box-containing (Pax) genes. Mutations in some of these Pax genes lead to severe developmental abnormalities. Zebrafish *pax[b]* (*pax[zf-b]*) is a member of the Pax gene family that is expressed in the presumptive posterior midbrain from the end of gastrulation and, at later stages, in other localized regions of the developing embryo¹. Here we show that injection of antibodies raised against the *pax[b]* protein causes a localized malformation at the midbrain–hindbrain boundary. *In situ* hybridizations demonstrate that antibody injection causes downregulation of *pax[b]* transcripts in the posterior midbrain and alteration of *wnt-1* and *eng-2* expression in this area. The data demonstrate an involvement of *pax[b]* in the formation of the midbrain–hindbrain junction.

Through homology to *Drosophila* paired box-containing segmentation genes, at least eight members of the vertebrate Pax gene family have been isolated¹⁻⁹. Each member of the family shows spatially and temporally restricted expression patterns during embryonic development. The paired box has recently been established as a DNA-binding element^{5,10,11}, thus Pax genes are likely to encode transcription factors. Mutations in several Pax genes lead to severe developmental abnormalities: (1) a mutation in the paired box of murine *Pax-1* leads to the *undulated* phenotype which exhibits distortions of the vertebral

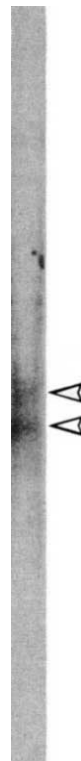


FIG. 1 The antibody to the *pax[b]* protein recognizes two polypeptides that closely correspond to the alternate splicing products of the *pax[b]* gene. The arrowheads demarcate the two antibody positive bands with estimated *M_r* of 43 and 45K. The two bands correlate well to the deduced *M_r* (42,028 and 43,747) of the two *pax[b]* splicing products¹.

METHODS. The rostral portion of the neural keels of 100 micro-dissected 14–16 h zebrafish embryos were homogenized in lysis buffer (20 mM Tris pH 7.8, 1 mM EDTA 1 mM PMSF, 0.5 mM DTT, 20 mM KCl) and aliquots of the homogenate were run on a 10% SDS-polyacrylamide gel. After transfer of the proteins to a nitrocellulose filter using a BioRad transfer system with a Tris/methanol/glycine buffer, strips of nitrocellulose were incubated in 100 mM Tris, 0.9% NaCl, 0.1% Tween-20, pH 7.5 (TTBS) for 1 h. Antibody incubations were done in TTBS overnight at 4 °C and control strips were similarly incubated in TTBS containing an equivalent amount of protein from normal rabbit serum. Second antibody and subsequent incubations were done according to the specifications of a Vectastain ABC Kit (Vector Laboratories) and colour was visualized with 3,3'-diaminobenzidine. The approximate molecular masses of the bands were calculated from the migration of known standards.

Received 11 August; accepted 15 September 1992.

1. Nurse, P., Thuriaux, P. & Nasmyth, K. *Molec. gen. Genet.* **146**, 167–178 (1976).
2. Nurse, P. *Nature* **344**, 503–508 (1990).
3. Minet, M., Nurse, P., Thuriaux, P. & Mitchison, J. M. *J. Bact.* **137**, 440–446 (1979).
4. Marks, J., Fankhauser, C. & Simanis, V. *J. Cell Sci.* **101**, 801–808 (1992).
5. Garnier, J., Osguthorpe, D. J. & Robson, B. *J. molec. Biol.* **120**, 97–120 (1978).
6. Côté, G. P. *Molec. cell Biochem.* **57**, 127–146 (1983).
7. Smillie, L. B. *Trends Biochem. Sci.* **4**, 151–155 (1979).
8. Kozak, M. *J. Cell Biol.* **108**, 229–241 (1989).
9. Liu, H. P. & Bretscher, A. *Cell* **57**, 233–242 (1989).
10. Goodwin, L. O., Lees-Miller, J. P., Leonard, M. A., Cheley, S. B. & Helfman, D. M. *J. Biol. Chem.* **266**, 8408–8415 (1991).
11. Liu, H. P. & Bretscher, A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 90–93 (1989).
12. Marks, J. & Hyams, J. S. *Eur. J. Cell Biol.* **39**, 27–32 (1985).
13. Alfa, C. E. & Hyams, J. S. *J. Cell Sci.* **96**, 71–77 (1990).
14. Jochová, J., Rupes, I. & Streiblová, E. *Cell Biol. Int. Rep.* **15**, 607–610 (1991).
15. Chant, J. & Herskowitz, I. *Cell* **65**, 1203–1212 (1991).
16. Schwob, E. & Martin, R. P. *Nature* **355**, 179–182 (1992).
17. Watts, F. Z., Shiels, G. & Orr, E. *EMBO J.* **6**, 3499–3505 (1987).
18. Schroeder, T. E. *Ann. N.Y. Acad. Sci.* **582**, 78–87 (1990).
19. Moreno, S., Klar, A. & Nurse, P. *Meth. Enzym.* **194**, 795–823 (1991).
20. Wright, A., Maundrell, K., Heyer, W.-D., Beach, D. & Nurse, P. *Plasmid* **15**, 156–158 (1986).
21. Tabor, S. & Richardson, C. C. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1074–1078 (1985).
22. Hagan, I. M. & Hyams, J. S. *J. Cell Sci.* **89**, 343–357 (1988).

ACKNOWLEDGEMENTS. We thank B. Wu, L. Fowke, J. Holmstrom and M. Whiting for assistance. P. Nurse for strains, P. G. Young for libraries, and H. E. Feilolter and P. G. Young for discussion. M.K.B. was supported by the University of Saskatchewan; D.M.H. was supported by the NIH and is an Established Investigator of the American Heart Association. Sequence presented in Fig. 1 has been submitted to GenBank, accession number L04126.

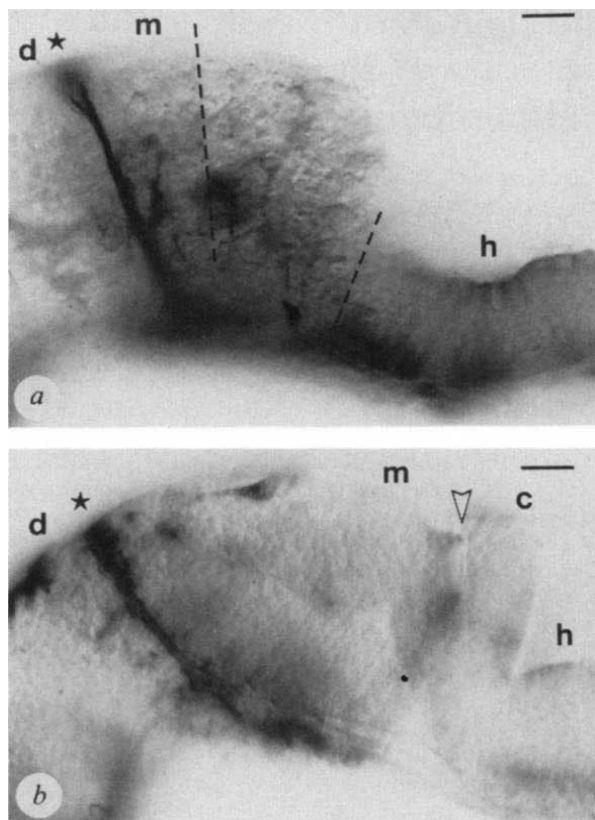


FIG. 3 *a*, Localized cell death in antibody-injected embryos at 27–28 h of development. Dorsal is up, rostral is to the left. The dashed lines demarcate the zone of dying cells. The embryos have been stained with an anti-acetylated α -tubulin antibody to visualize axon tracts. Dying cells were identified by their opacity and highly refractile appearance³². The extent of cell death varied between embryos and between the two sides of the neural tube in individual embryos. The embryo illustrated demonstrates the most marked cell death that was observed. Note that the affected cells extend rostrally and caudally beyond the region that fails to express *pax[b]*. In phenotypically abnormal brains, there was some disorganized differentiation of neurons caudal to the posterior commissure. The posterior commissure itself (marked by a star) was unaffected. *b*, An embryo with normal phenotype. *d*, Diencephalon; *c*, cerebellum; *h*, hindbrain; *m*, midbrain. Scale bars, 25 μ m.

column¹⁰; (2) a mutation in the *Pax-6* gene leads to the semi-dominant *Small eye* phenotype in mouse¹² and a mutation in the putative human homologue of this gene may be responsible for the aniridia disorder¹³; (3) mutations in the putative human homologue to *Pax-3* (*HuP2*) occasionally lead to hearing loss associated with Waardenburg's syndrome^{14,15}, whereas in mouse a deletion in *Pax-3* causes the *splotch* phenotype which is associated with spina bifida, exencephaly, a tail flexion defect and deficiencies in neural crest cell derivatives¹⁷.

To address the role of *pax[b]* protein (which has highest homology to murine *Pax-2*) during zebrafish development, an antibody specific for the putative transcription activation site of the *pax[b]* protein^{1,16} was injected into fertilized zebrafish eggs. The antibody exclusively stains nuclei of cells that also express the *pax[b]* gene¹⁶. Furthermore, no crossreactivity was found on western blots as the two unique bands (Fig. 1) correlate well with the predicted relative molecular mass of the two *pax[b]* splicing products¹. Injections of the antibody did not significantly increase mortality rates compared to controls and until 22 h of development no obvious morphological defects were observed. At about this time, when the furrow that separates the midbrain from the hindbrain becomes apparent, 50–60% of the injected embryos showed abnormal development of this region. The nature of the abnormality varied with the amount of antibody injected. Smaller injections led to a hypertrophy of

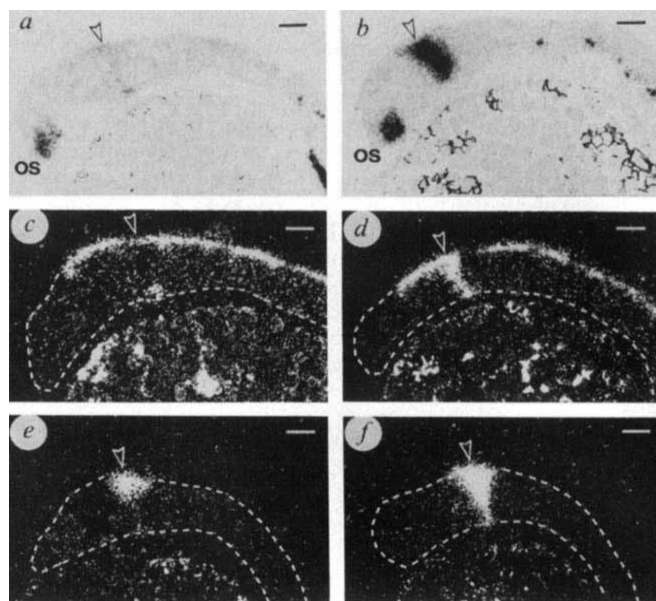


FIG. 4 *In situ* hybridizations on tissue sections of 14-h embryos that were microinjected with the antibody against *pax[b]* protein (*a, c, e*), and wild-type embryos (*b, d, f*). Bright-field (*a, b*) and dark-field (*c–f*) photographs are shown on sagittal sections (the section in *e* is slightly parasagittal). Rostral is to the left. The following probes were used: *a, b*, 2,409-bp *pax[b]* cDNA¹; *c, d*, 503-bp *Bam*HI–*Hind*III 3' *wnt-1* genomic fragment²⁵; *e, f*, 1,116-bp *eng-2* cDNA. The region of the presumptive posterior midbrain is marked by an arrowhead. *os*, Optic stalk.

METHODS. *In situ* hybridizations were done as described in ref. 1. Slides for tissue sections were treated with 3-aminopropyl-triethoxysilane rather than gelatinized. Sections were exposed for 14 days. Scale bars, 50 μ m.

the midbrain–hindbrain boundary region (eight cases; 12–15% of the injected embryos), whereas embryos that were injected with a high dose of the antibody failed to form the midbrain–hindbrain boundary (32 cases; 50–60% of the injected embryos). Instead, the embryos possessed a neuroepithelium that gradually tapered towards the diencephalon (Fig. 2). Other regions of *pax[b]* expression did not show major morphological changes, but we cannot rule out that subtle changes have escaped examination. At 27–28 h of development, localized cell death was seen in antibody-injected embryos in the region that would normally form the posterior portion of the midbrain and cerebellum (Fig. 3). Control injections with preimmune serum produced normal embryos.

To elucidate molecular changes preceding the morphological abnormalities, *in situ* hybridizations were done on tissue sections of injected embryos at 14 h of development. In wild-type embryos, *pax[b]* transcripts are first detected at 9 h (at the end of gastrulation) in tissue that will later form the caudal midbrain¹. At 14 h of development, additional *pax[b]* expression is detected in the optic stalk, the otic placode, the nephritic primordium and in presumptive commissural interneurons along the hindbrain and the spinal cord^{1,16}. At 14 h of development, *pax[b]* expression in the posterior midbrain of uninjected embryos was prominent compared to *pax[b]* expression in the stalk (Fig. 4*b*). In contrast, *pax[b]* expression in the posterior midbrain of injected embryos was dramatically reduced, whereas the amount of *pax[b]* transcript in the stalk did not show major changes (Fig. 4*a*).

To analyse whether other putative developmental control genes expressed at the midbrain–hindbrain boundary were affected by antibody injections, we examined the expression of the zebrafish homologues of murine *Wnt-1* and *En-2*, genes that are important for the normal formation of this region in mice^{18–22}. *Wnt-1*, a gene that encodes a signalling peptide^{23,24}, is expressed in the neural keel of zebrafish embryos at 14 h of

FIG. 2 Injection of the antibody to *pax[b]* protein leads to abnormal development of the midbrain-hindbrain boundary. *a*, Phenotype created by injecting fertilized zebrafish eggs with antibody compared to *b*, the phenotype of a normal embryo at 24 h of development. Dorsal views of the embryos in which rostral is to the top. The large arrowhead marks the position of the furrow that separates midbrain and hindbrain in normal embryos (*b*) and the position where the furrow is missing in antibody-injected embryos (*a*). The small arrowhead indicates the unaffected rhombomere, ro2, as marked by axon staining with an anti-acetylated α -tubulin antibody (kindly provided by G. Piperno). f, Forebrain; h, hindbrain; m, midbrain. Scale bars, 25 μ m.

METHODS. Fertilized zebrafish eggs were injected on an agarose ramp through the intact chorion and into the yolk. Injections of ~ 5 nl of the antibody ($40 \mu\text{g } \mu\text{l}^{-1}$) generated embryos that failed to form the furrow between midbrain and hindbrain in 50–60% of cases, whereas injection of ~ 2 nl of antibody led to hypertrophy of the furrow region in 12–15% of cases. Embryos were raised at 28.5 °C. The antibody was present in the embryos until at least 28 h of development as demonstrated by staining embryos with fluorescently labelled secondary antibodies to the injected primary antibody (not shown). The staining with the anti-acetylated α -tubulin antibody was done using standard procedures.



development, extending from the region of the presumptive epiphysis to the tail (excluding the presumptive cerebellum), and in a transverse stripe in the posterior midbrain (ref. 25 and Fig. 4d). In injected embryos the dorsal portion of the *wnt-1* expression appeared to be unaltered, but the expression in the transverse midbrain stripe was strongly reduced (Fig. 4c). In zebrafish (for mouse, see ref. 26), transcripts of *engrailed* genes are confined to the posterior midbrain and the presumptive cerebellum at 14 h of development (Fig. 4f). *In situ* hybridizations on injected embryos showed that, at this developmental stage, transcripts of *eng-2* were only detected in the dorsal portion of the neural keel at the midbrain-hindbrain boundary (Fig. 4e).

Our results provide evidence that the *pax[b]* gene is a key regulatory element in the formation of the posterior midbrain. The alteration in the level of *pax[b]* transcripts in antibody-injected embryos suggests that a functional *pax[b]* protein is required (directly or indirectly) for the expression of its own gene in the posterior midbrain. Such autoregulatory pathways are known for *Drosophila* segmentation genes²⁷. A consequence of the reduced *pax[b]* expression in the posterior midbrain (and the interference of the antibody with the *pax[b]* protein) seems to be a downregulation of two other genes in this region, *wnt-1* (*Drosophila* homologue: *wingless*) and *eng-2* (*Drosophila* homologue: *engrailed*). This observation may explain the resemblance between the '*pax[b]*' phenotype and the *Wnt-1*[−]/*Wnt-1*[−] phenotype²², and suggests a regulatory network between *pax[b]*, *wnt-1* and *eng-2* in vertebrates.

In *Drosophila* the segmentation genes *paired* and *gooseberry* are required for a proper *wingless* and *engrailed* expression^{28–30}. Interestingly, both genes contain a paired box³¹ and share homology with *pax[b]*¹. Thus comparable networks of interacting factors may have evolved in *Drosophila* and vertebrates. □

Received 2 April; accepted 24 September 1992.

1. Krauss, S., Johansen, T., Korzh, V. & Fjose, A. *Development* **113**, 1193–1206 (1991).
2. Deutsch, U., Dressler, G. R. & Gruss, P. *Cell* **53**, 617–625 (1988).
3. Dressler, G. R., Deutsch, U., Chowdhury, K., Nornes, H. O. & Gruss, P. *Development* **109**, 787–795 (1990).
4. Plachov, D. et al. *Development* **110**, 643–651 (1990).
5. Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R., & Gruss, P. *EMBO J.* **10**, 1135–1147 (1991).
6. Jostes, B., Walther, C., & Gruss, P. *Mech. Dev.* **33**, 27–38 (1991).
7. Krauss, S. et al. *EMBO J.* **10**, 3609–3619 (1991).
8. Walther, C. & Gruss, P. *Development* **113**, 1435–1449 (1991).
9. Walther, C. et al. *Genomics* **11**, 424–434 (1991).
10. Chalepakis, G. et al. *Cell* **66**, 873–884 (1991).
11. Treisman, J., Harris, E. & Desplan, C. *Genes Dev.* **5**, 594–604 (1991).
12. Hill, R. E. et al. *Nature* **354**, 522–525 (1991).
13. Ton, C. C. T. et al. *Cell* **67**, 1059–1074 (1991).
14. Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silva, E. O. & Milunsky, A. *Nature* **355**, 637–638 (1992).
15. Tassabehji, M. et al. *Nature* **355**, 635–636 (1992).
16. Mikkola, J. et al. *J. Neurobiol.* **23**, 933–946 (1992).
17. Epstein, D. J., Vekemans, M. & Gros, P. *Cell* **67**, 767–774 (1991).
18. McMahon, A. P. & Bradley, A. *Cell* **62**, 1073–1085 (1990).
19. Thomas, K. R. & Capocchi, M. R. *Nature* **346**, 847–850 (1990).
20. Joyner, A. L., Herrup, K., Auerbach, B. A., Davis, C. A. & Rossant, J. *Science* **251**, 1239–1243 (1991).
21. Thomas, K. R., Musci, T. S., Neumann, P. E. & Capocchi, M. *Cell* **67**, 969–976 (1991).
22. McMahon, A. P., Joyner, A. L., Bradley, A. & McMahon, J. A. *Cell* **69**, 581–595 (1992).
23. Nüsse, R. & Varmus, H. E. *Cell* **31**, 99–109 (1982).
24. Papkoff, J., Brown, A. M. C. & Varmus, H. E. *Molec. cell Biol.* **7**, 3978–3984 (1987).
25. Molven, A., Njåstad, P. R. & Fjose, A. *EMBO J.* **10**, 799–807 (1991).
26. Joyner, A. L., Kornberg, T., Coleman, K. G., Cox, D. R. & Martin, G. R. *Cell* **43**, 29–37 (1985).
27. Hiromi, Y. & Gehring, W. J. *Cell* **50**, 963–974 (1987).
28. DiNardo, S. & O'Farrell, P. H. *Genes Dev.* **1**, 1212–1225 (1987).
29. Ingham, P. W., Baker, N. E. & Martinez-Arias, A. *Nature* **331**, 73–75 (1988).
30. Hidalgo, A. & Ingham, P. *Development* **110**, 291–301 (1990).
31. Bopp, D., Burri, M., Baumgartner, S., Frigerio, G. & Noll, M. *Cell* **47**, 1033–1049 (1986).
32. Grunwald, D. J., Kimmel, C. B., Westerfield, M., Walker, C. & Streisinger, G. *Dev. Biol.* **126**, 115–128 (1988).

ACKNOWLEDGEMENTS. We thank P. Ingham and R. Patient for reading the manuscript, A. Lumsden and C. Wylie for discussions, T. Eskeland, T. Johansen, J. Kuwada and J. Pizze for support at various stages of this project, I. Mikkola for the work that led to the generation of the antibody which was done in the laboratory of A. Fjose, whom we thank for his cooperation, and A. Molven and N. Patel for *wnt-1* and *eng-2* probes, respectively. This work was supported financially in part by an SERC programme grant to N.H., M.M. and S.W.W. S.K. is a DNK fellow. S.W.W. is a SERC Advanced Research Fellow. The initial experiments were done at the University of Michigan, Ann Arbor.

Insights from the classic work of Luria and Delbrück on bacterial mutation rates may be applied to tracking human gene defects prevalent in isolated populations.

It takes a certain panache to equate the exponential growth of a bacterial suspension in a conical flask to the more sedate expansion of a human population the size of, say, Finland, and furthermore to use the analogy to deduce the position of the gene for an uncharacterized hereditary trait. But in this month's *Nature Genetics*, Lander, de la Chapelle and colleagues suggest² that the classic experiments of Luria and Delbrück¹, widely recognized as a landmark in the history of genetics research, may take on new significance in providing a means to pinpoint the genes responsible for hereditary traits in isolated populations².

The general problem facing de la Chapelle and coworkers is how to progress from a linked genetic marker for a hereditary disorder to the disease locus in an efficient manner — the process of positional cloning³. They are not the only ones: following the successful application of chromosomal linkage analysis to map a gene, researchers are often faced with the daunting task of navigating across a million or more base pairs of DNA — how then to prove they have identified the correct gene when there may be dozens of other coding sequences close by? Often the identification of chromosomal deletions and rearrangements in affected patients yields a relatively small, critical region of DNA; but in at least two notable examples, cystic fibrosis (CF) and Huntington's disease, this was not the case.

One technique that is often applied to narrow the gap is linkage disequilibrium, which relies on the fact that markers in close proximity to the disease mutation are likely to remain associated with the disease chromosome — evidence of the genetic background on which the original mutation occurred. Such an approach was spectacularly successful in identifying the CF gene in 1989 — there is good evidence that the most common CF mutation, $\Delta F508$ (deletion of a phenylalanine at residue 508), arose only once

on a single haplotype — but so far has not been as useful in other settings. However, de la Chapelle and colleagues² now point out that in certain situations, such as when a population is isolated and contains little immigration, and when the disease is relatively frequent, linkage disequilibrium can prove to be an especially powerful technique.

Such circumstances apply to diastrophic dysplasia (DTD), one of more than 80 distinct osteochondrodysplasias, which is characterized by short stature, dysplasia and calcification of the joints, spinal anomalies, 'hitch-hiker' thumbs, clubbed feet and, often, cleft palate. Although found in most populations, DTD is particularly common in Finland, with an allele frequency of 0.8 per cent in a country with a population of 5 million. The DTD gene was finally mapped to the long arm of chromosome 5 in 1990 (ref. 4), but even in the wake of considerable refinement of the localization^{2,5}, the DTD locus could still only be assigned to an area of about 1,500 kilobases (kb).

Although a good candidate for linkage disequilibrium approaches, it is the application of the 1943 study of Salvador Luria and Max Delbrück on the mutation rates of bacterial populations¹ which is especially intriguing. Luria and Delbrück showed conclusively that *Escherichia coli* acquire mutations (such as resistance to bacteriophage) spontaneously, rather than acquiring them by contact with the viruses, for example, as some had suggested. A comparison of several separate cultures of *E. coli* exposed to virus revealed large differences in the number of resistant colonies, depending on when the mutation had occurred during culture, as predicted by a theoretical assessment of the mutation rate.

Hastbäck *et al.*² argue that, to a first approximation, this situation resembles the spread of DTD in the Finnish population: a single mutation spreading in an isolated population, but which occurred sufficiently long ago that recombination around the DTD locus has shuffled all markers with the likely exception of those closest to the DTD gene. A single ancestral DTD mutation is likely as 95% patients bear the same 1-1 haplotype, defined by two polymorphisms at a flanking locus (colony-stimulating factor-1 receptor, *CSF1R*).

By assuming that all DTD alleles in Finland arose from a common ancestor, and all recombinations (5%) between DTD and *CSF1R* removed the common 1-1 haplotype, the authors conclude — applying the equation formulated by Luria and Delbrück¹ ($1 - e^{-g\theta} = \pi$, where g = an estimated 100 generations, θ = recombination frequency, and π = proportion of recombinant chromosomes (0.05)) — that DTD must lie within about 0.06 centimorgans of *CSF1R*, roughly equivalent to 60 kb. And if a proportion of the 5% DTD patients who have a haplotype other than 1-1 are due to new mutations, this distance will be even smaller.

The advantage of the authors' analysis is that they extract genetic information from the cumulative genetic history of Finland's population (essentially founded about 2,000 years ago). During this time recombination has confined the region of disequilibrium, but few new mutations are likely to have arisen. In the long term, it will be fascinating to see if such an approach can be extended to other disorders and/or other isolated populations. These might include well-known groups like the Amish (often the subjects of genetic studies), but their relative youth may not permit such tightly focused regions of disequilibrium.

More immediately, the specific prediction for DTD is that the gene has been narrowed to just 60 kb, representing a 25-fold increase in resolution over conventional techniques. Hastbäck *et al.* detect a recombination event in one of the introns of *CSF1R* that places DTD proximal to this gene. But immediately upstream of *CSF1R* is the gene for the platelet-derived growth-factor receptor β -subunit — neither is an obvious candidate for the DTD gene. Future studies should soon uncover the position and identity of the DTD gene, information that may well indicate that we can still learn a thing or two from the humble bacterium.

Kevin Davies

Kevin Davies is editor of Nature Genetics.

Also in this month's *Nature Genetics*: thousands more expressed genes from liver and brain tissues sequenced and characterized; the search for new genes containing trinucleotide-repeat sequences; expression studies of *XIST* RNA in human and mouse germ cells, and CFTR in airway submucosal glands in normal and CF individuals.

1. Luria, S. & Delbrück, M. *Genetics* **28**, 491–511 (1943).

2. Hastbäck, J. *et al. Nature Genet.* **2**, 204–211 (1992).

3. Collins, F. *Nature Genet.* **1**, 3–6 (1992).

4. Hastbäck, J., Kaitila, I., Sistonen, P. & de la Chapelle, A. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8056–8059 (1990).

5. Hastbäck, J. *et al. Genomics* **11**, 968–973 (1991).